

Philosophical Studies Series

Bernardo Ferro

Masters, Slaves and Philosophers

Plato, Hegel and Nietzsche on Freedom
and the Pursuit of Knowledge



Springer

Philosophical Studies Series

Volume 149

Editor-in-Chief

Mariarosaria Taddeo, Oxford Internet Institute, University of Oxford, Oxford, UK

Editorial Board Member

Patrick Allo, Vrije Universiteit Brussel, Brussel, Belgium

Advisory Editors

Lynne Baker, Department of Philosophy, University of Massachusetts, Amherst, USA

Stewart Cohen, Arizona State University, Tempe, AZ, USA

Radu Bogdan, Department of Philosophy, Tulane University, New Orleans, LA, USA

Marian David, Karl-Franzens-Universität, Graz, Austria

John Fischer, University of California, Riverside, Riverside, CA, USA

Keith Lehrer, University Of Arizona, Tucson, AZ, USA

Denise Meyerson, Macquarie University, Sydney, Australia

Francois Recanati, Ecole Normale Supérieure, Institut Jean Nicod, Paris, France

Mark Sainsbury, University of Texas at Austin, Austin, TX, USA

Barry Smith, State University of New York at Buffalo, Buffalo, NY, USA

Nicholas Smith, Department of Philosophy, Lewis and Clark College, Portland, OR, USA

Linda Zagzebski, Department of Philosophy, University of Oklahoma, Norman, OK, USA

Philosophical Studies Series aims to provide a forum for the best current research in contemporary philosophy broadly conceived, its methodologies, and applications. Since Wilfrid Sellars and Keith Lehrer founded the series in 1974, the book series has welcomed a wide variety of different approaches, and every effort is made to maintain this pluralism, not for its own sake, but in order to represent the many fruitful and illuminating ways of addressing philosophical questions and investigating related applications and disciplines.

The book series is interested in classical topics of all branches of philosophy including, but not limited to:

- Ethics
- Epistemology
- Logic
- Philosophy of language
- Philosophy of logic
- Philosophy of mind
- Philosophy of religion
- Philosophy of science

Special attention is paid to studies that focus on:

- the interplay of empirical and philosophical viewpoints
- the implications and consequences of conceptual phenomena for research as well as for society
- philosophies of specific sciences, such as philosophy of biology, philosophy of chemistry, philosophy of computer science, philosophy of information, philosophy of neuroscience, philosophy of physics, or philosophy of technology; and
- contributions to the formal (logical, set-theoretical, mathematical, information-theoretical, decision-theoretical, etc.) methodology of sciences.

Likewise, the applications of conceptual and methodological investigations to applied sciences as well as social and technological phenomena are strongly encouraged.

Philosophical Studies Series welcomes historically informed research, but privileges philosophical theories and the discussion of contemporary issues rather than purely scholarly investigations into the history of ideas or authors. Besides monographs, *Philosophical Studies Series* publishes thematically unified anthologies, selected papers from relevant conferences, and edited volumes with a well-defined topical focus inside the aim and scope of the book series. The contributions in the volumes are expected to be focused and structurally organized in accordance with the central theme(s), and are tied together by an editorial introduction. Volumes are completed by extensive bibliographies.

The series discourages the submission of manuscripts that contain reprints of previous published material and/or manuscripts that are below 160 pages/88,000 words.

For inquiries and submission of proposals authors can contact the editor-in-chief Mariarosaria Taddeo via: mariarosaria.taddeo@oii.ox.ac.uk

More information about this series at <https://link.springer.com/bookseries/6459>

Bernardo Ferro

Masters, Slaves and Philosophers

Plato, Hegel and Nietzsche on Freedom and
the Pursuit of Knowledge



Springer

Bernardo Ferro
Institute for Philosophical Studies
University of Coimbra
Coimbra, Portugal

ISSN 0921-8599 ISSN 2542-8349 (electronic)
Philosophical Studies Series
ISBN 978-3-030-90404-3 ISBN 978-3-030-90405-0 (eBook)
<https://doi.org/10.1007/978-3-030-90405-0>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Abbreviations and Translations

All translations of Plato's works are taken from the Loeb edition, published by Harvard University Press, with occasional modifications. The dialogues are identified by the following abbreviations, followed by the relevant Stephanus numbers.

<i>Ap</i>	<i>The Apology</i>	<i>Phd</i>	<i>Phaedo</i>
<i>Cra</i>	<i>Cratylus</i>	<i>Phdr</i>	<i>Phaedrus</i>
<i>Cri</i>	<i>Crito</i>	<i>Phi</i>	<i>Philebus</i>
<i>Eu</i>	<i>Euthyphro</i>	<i>Pr</i>	<i>Protagoras</i>
<i>Go</i>	<i>Gorgias</i>	<i>Re</i>	<i>The Republic</i>
<i>La</i>	<i>Laws</i>	<i>So</i>	<i>The Sophist</i>
<i>Ly</i>	<i>Lysis</i>	<i>St</i>	<i>The Statesman</i>
<i>Me</i>	<i>Meno</i>	<i>Sy</i>	<i>Symposium</i>
<i>Pa</i>	<i>Paemenides</i>	<i>Th</i>	<i>Theaetetus</i>

As regards Hegel's works, reference is given first to the English translation, indicated by the abbreviations listed below, followed by a slash and the indication of the German edition. All German references are taken from *Werke in zwanzig Bänden* (eds. E. Moldenhauer and K. M. Michel. Frankfurt: Suhrkamp, 1986), abbreviated as 'HW' and followed by the volume and page numbers. Where no English edition is indicated, the translation is mine.

- DFS *The Difference Between Fichte's and Schelling's System of Philosophy* (Albany: State of New York University Press, 1977)
- EL *Encyclopedia of the Philosophical Sciences in Basic Outline. Part I: Science of Logic* (Cambridge: Cambridge University Press, 2010)
- ETW *Early Theological Writings* (Pennsylvania: University of Pennsylvania Press, 1971)
- FK *Faith and Knowledge* (Albany: State University of New York Press, 1977)
- LHP *Lectures on the History of Philosophy* (Lincoln: Nebraska University Press, 1995)
- PM *Philosophy of Mind* (Oxford: Oxford University Press, 2007)
- PN *Philosophy of Nature* (Oxford: Oxford University Press, 1970)
- PR *Outlines of the Philosophy of Right* (Oxford: Oxford University Press, 2008)
- PS *Phenomenology of Spirit* (Oxford: Oxford University Press, 1970)
- SL *Science of Logic* (Cambridge: Cambridge University Press, 2010)

As regards Nietzsche's works, reference is given first to the English translation, indicated by the abbreviations listed below, followed by a slash and the indication of the German edition. All German references are taken from the *Kritische Studienausgabe* (eds. G. Colli and M. Montinari. Berlin: Walter de Gruyter, 1988), abbreviated as 'KSA' and followed by the volume and page numbers. Where no English edition is indicated, the translation is mine.

- A *The Antichrist, Ecce Homo, Twilight of the Idols* (Cambridge: Cambridge University Press, 2005)
- BGE *Beyond Good and Evil* (Cambridge: Cambridge University Press, 2002)
- BT *The Birth of Tragedy and Other Writings* (Cambridge: Cambridge University Press, 1999)
- D *Daybreak: Thoughts on the Prejudices of Morality* (Cambridge: Cambridge University Press, 1999)
- GM *On the Genealogy of Morality* (Cambridge: Cambridge University Press, 2006)
- GS *The Gay Science. With a Prelude in German Rhymes and an Appendix of Songs* (Cambridge: Cambridge University Press, 2001)
- HAH *Human, All Too Human: A Book for Free Spirits* (Cambridge: Cambridge University Press, 1996)
- TSZ *Thus Spoke Zarathustra: A Book for All and None* (Cambridge: Cambridge University Press, 2006)

As regards the works by other authors, where no English edition is indicated the translation is mine. Unless otherwise noted, all quotations retain their original emphases.

Acknowledgements

This book is a substantially revised version of my first doctoral dissertation, presented in 2017 at the New University of Lisbon. What started out as a study on Hegel's engagement with Plato and the Platonic tradition ended up becoming a much wider historical survey, centred on three of the most important, fascinating, but also difficult thinkers of the history of philosophy. In what follows, I revisit the works of Plato, Hegel and Nietzsche in order to draw a brief conceptual genealogy of one of philosophy's oldest and most significant debates: the question of whether, how and to what extent human freedom depends on the acquisition of knowledge.

Before getting started, I would like to thank all of those who contributed to the successful completion of this project. First of all, a special mention is due to my supervisor, Mário Jorge de Carvalho, to whom I owe most of my philosophical training. His brilliant lectures on Plato's dialogues and Hegel's *Phenomenology of Spirit*, given at the New University Lisbon, provided the inspiration for this work, and his invaluable advice and criticism have helped me find my way in a complex and often confusing philosophical labyrinth.

During most of my PhD, my research was funded by the Portuguese Foundation for Science and Technology (FCT), through my membership of the Institute for Philosophical Studies (IEF), and by the German Academic Exchange Service (DAAD), which allowed me to spend 1 year at Humboldt University in Berlin. I am grateful to Tobias Rosefeldt for supporting my application and letting me take part in his weekly research seminar, to Bernhard Thöle for his interest in my work and especially to Franz Knappik for his helpful comments on an earlier version of the manuscript. I am also particularly indebted to Karin de Boer for her unfailing generosity and her shrewd criticism during the preparation of this book.

Finally, I would also like to thank Max Marcuzzi, Jacinto Rivera de Rosales, Nuno Ferro, Martim Enes Dias, Lucas Manarte and Hélder Telo for their philosophical and literary input, as well as my family and friends for their support and encouragement during the last years. I dedicate this book to all of them.

Contents

Abbreviations and Translations	v
Acknowledgements	vii
1 Introduction	1
1.1 A Brief Anatomy of Freedom	3
1.2 Freedom and Education	11
1.3 Plato, Hegel and Nietzsche	20
1.3.1 Plato's Philosophical Project	21
1.3.2 Hegel's Phenomenological Project	26
1.3.3 Nietzsche's Metacognitive Project	30
1.4 Historical Affinities	33
1.5 Research Background	37
Part I Plato's Philosophical Project	
2 Plato's Idea of Truth	47
2.1 Truth, Opinion and Doxa	48
2.2 Self-Confidence and Self-Interest	58
2.3 Body and Soul	65
3 Masters, Slaves and Philosophers	75
3.1 Rhetors and Philosophers	76
3.2 Socrates Versus Callicles	82
3.3 Masters and Slaves	87
4 Chains and Shadows	97
4.1 Inside the Cave	99
4.2 Towards the Light	107
4.3 The Self as Shadow	113

5	Philosophy and Deliverance	117
5.1	The Need for Deliverance	119
5.2	The Possibility of Deliverance	123
5.3	Plato and Hegel	128

Part II Hegel's Phenomenological Project

6	Hegel's Phenomenology of Spirit	135
6.1	Truth and Appearance	137
6.2	Experience as Science	145
6.3	Further Hermeneutical Issues	152
7	Knowledge, Power and Freedom	159
7.1	Selfhood and Otherness	160
7.2	Selfhood as Recognition	169
7.3	Masters, Slaves and Philosophers	176
7.4	The Unhappy Consciousness	187
8	Towards Freedom	195
8.1	Reason and Spirit	196
8.2	The Trials of Spirit	206
8.3	Tragedy and Comedy	214
9	Beyond the Phenomenology	219
9.1	Final Revelations	221
9.2	Science and the System	225
9.3	Hegel and Nietzsche	231

Part III Nietzsche's Metacognitive Project

10	Nietzsche's Idea of Truth	237
10.1	The Will to Truth	240
10.2	Truth and Metaphor	246
10.3	Science and Wisdom	253
11	Philosophy in Chains	263
11.1	The Socratic Turn	264
11.2	Athens and Jerusalem	268
11.3	Masters and Slaves	274
11.4	Asceticism and Nihilism	281
12	Philosophy Unchained	287
12.1	The Will to Power	288
12.2	Return to Callicles	293
13	Conclusions	301
	Bibliography	305
	Index	311

Did you not see the helpless infirmity,
no better than a dream, in which
the blind generation of men is shackled?

AESCHYLUS (?), *Prometheus Bound*

Chapter 1

Introduction



Abstract The introduction offers a preliminary discussion of the relationship between freedom, power and knowledge, its philosophical relevance and its conceptual and historical implications. Starting from a general definition of freedom, I concentrate on freedom's cognitive dimension. My main focus is the extent to which a free life requires a clear understanding of reality—i.e. the extent to which knowledge is empowering or *liberating*; and, conversely, the extent to which ignorance is debilitating or *enslaving*. I argue that this issue was first raised in Plato's dialogues, where human ignorance is portrayed as a form of enslavement, and went on to become an important philosophical theme.

After introducing Plato's position, I highlight its unique influence on the work of two modern authors, Hegel and Nietzsche, who explicitly linked the notions of ignorance and truth to those of bondage and freedom—or slavery and mastery. I claim that these two philosophical projects represent, in many ways, the most extreme responses to Plato's founding idea. While the dialogues stressed the need to re-educate, and thereby liberate human consciousness, Hegel and Nietzsche went further and sought to fulfil this ambitious task, but their solutions point in very different directions: whereas Hegel believes in the possibility of a true, and hence truly free mode of cognition, Nietzsche calls into question the very notion of truth, as well as the freedom it is supposed to offer.

Keywords Freedom · Power · Knowledge · Truth · Plato · Hegel · Nietzsche

Out of all the issues that make up the history of Western philosophy, that of human freedom—of its reality, nature and scope—is undoubtedly one of the most important. However, as is often the case with philosophical issues, it rarely comes to mind in the course of most people's lives. Unless one's freedom is somehow called into question and must be reasserted or redefined; or unless some practical need leads one to focus on or wonder about someone else's freedom; or unless one is a professional philosopher, or barrister, or magistrate, and makes a living out of dealing with such matters, the questions of whether, how and to what extent human beings are free are hardly ever asked. And this is so because the aforementioned scenarios are not the

rule, but the exception. Most people's lives are spent dealing with more mundane issues and more immediate concerns. Although the importance of freedom is generally acknowledged, its exact definition is usually seen as a strictly theoretical problem; and even when one appreciates its practical implications, the fact remains that a thorough answer to the questions of whether, how and to what extent human beings are free is simply not needed when, say, moving around, speaking or pouring a cup of coffee.

But even though freedom as such is not often defined or discussed, it is nonetheless an essential component of most people's lives. For although one does not need to wonder about the philosophical meaning of freedom in order to move around, or speak, or pour a cup of coffee, one does all of these things in the assumption that one is free to do them—or not to do them, or to do different things. Precisely because it is so familiar, the ability to think and act freely is usually presupposed and not explicitly acknowledged. Just like breathing or having a body, it ranks among the most basic requirements of human experience and provides the ground on which most thoughts and actions usually rest. In other words, freedom is part of the *medium* or the *atmosphere* in which human life unfolds, which means not only that most human beliefs are silently supported by the idea of freedom, but also that this idea is rarely, if ever, truly problematized. Whenever freedom is questioned, criticized or found wanting, one tends to perceive one's lack of freedom as a contingent loss, which calls for concrete action and specific forms of liberation. It is at best an ethical or political problem, leading to disagreement and strife, to war and revolution. Paradoxically, however, the freedom to arrive at this diagnosis—the freedom that allows one to identify and expose one's occasional lack of freedom—is still taken for granted. Only the surface of the issue has been stirred, while its inner core remains untouched.

A more serious investigation, however, is bound to produce different results. If freedom is a basic ingredient of human life, the question of whether we are as free as we take ourselves to be must have a direct bearing not only on this or that particular judgement or decision, but on our entire understanding of reality. Since our general conviction regarding the freedom of our thoughts and actions provides the basis for the usual running of our lives, the abandonment of this conviction, or the endorsement of a different conviction, is sure to have profound practical consequences. Thus construed, the debate concerning human freedom acquires a deeper ontological significance, which calls for a re-examination of what it means to be an independent, self-conscious individual. It may be, of course, that this enquiry is destined to lead nowhere. If personal freedom, like breathing, is indeed a vital necessity, to question its actuality may prove as futile as picturing one's own death—or as absurd as Baron Münchhausen's ability to pull himself up by his own hair. Even so, there is still a long road separating our automatic belief in freedom from the actual truthfulness of that belief. And this road might reveal, if nothing else, the extraordinary complexity of the notion of freedom. In particular, it might show how different definitions of freedom correspond to different outlooks on reality, and how the latter imply radically different interpretations of what is at stake in human life.

1.1 A Brief Anatomy of Freedom

What is the origin of the confidence we feel with regard to our own freedom? How truthful is the assumption that we are free and how wide is its scope? The answers to these questions are far from simple, and their complexity derives, to a great extent, from the ambiguity of our usual definition of freedom.

To begin with, it is easy to see that the notion of freedom is not simple or univocal. It can come up in different contexts with very different meanings. Accordingly, our usual conviction regarding our own freedom is neither simple nor univocal. Like most immediate assumptions, it can be interpreted in very different ways. More importantly, though, this assumption conceals a great deal of ignorance. Although we usually fail to notice it, it really amounts to a series of overlapping assumptions, whose exact number, meaning and articulation are anything but clear. Therefore, to understand what is at stake in our usual belief in personal freedom, we must start by dissecting its different components and determining how they relate to one another.

Let us consider a few verses from one of Schiller's late poems, partly dedicated to the issue of freedom. John Sigerson's skilful translation attempts to recreate the style and rhyming of the original verses:

Man was created free—*is* free,
 E'en though he were born in shackles.
 Do not be deceived by the rabble's bray
 Or idiots' abusive cackles.
 Before the slave, when his chains he doth break,
 Before the man who's free, O do not quake!¹

Man is free—even though he were born in shackles. In these well-known verses, Schiller suggests that freedom is an essential feature of humanity, whereas bondage or servitude is not. He distinguishes between what might be called a primary or essential kind of freedom, inherent to the very definition of humanity, and a secondary or contingent one, which may or may not be actualized in the course of human life. Even though all human beings are naturally free, some live freely while others do not. And even when personal freedom seems non-existent, as in the case of life-long slavery and oppression, bondage and servitude are still somehow accidental predicates, forced upon us by contingent circumstances.

This idea is reiterated by Schiller's vigorous defence of human freedom against the 'bray' and the 'abusive cackles' of the 'rabble', that is, against the vain utterings of those who seek to undermine personal freedom. And it is further developed throughout the poem, where Schiller argues that if humans forget the true meaning

¹Schiller, *Friedrich Schiller*, vol. 1, 335

of the word freedom, or lose faith in it, they are ‘deprived of all their value’.² Now this claim could be read in a less definitive way. It might simply mean that those who deny human beings their right to freedom are cruel or unfair, and that men and women should be more highly valued. This is the kind of humanist claim that is generally championed in modern Western societies and that most reasonable people are ready to endorse. But Schiller’s point is more ambitious. In his eyes, those who deny human beings their right to freedom are fundamentally mistaken, as they fail to acknowledge their true value. In other words, they fail to recognize that human beings are *intrinsically* free, even if they are born or live in shackles. In this second reading, what is at stake is not so much human beings’ actual life as their essence or nature. To question a man’s freedom amounts to nothing less than to question what it means to be human.

This second claim is not as straightforward as the first. If we consider our daily lives, there seems to be no doubt as to the limited scope of our freedom. Even if we were not born or live in shackles, in the common sense of the word, it seems fairly obvious that we cannot do whatever we want whenever we want to and that many of our choices are instinctive or involuntary. Indeed, hardly a day goes by without our being reminded of our impotence or faced with situations whose meaning eludes us and whose outcome escapes our control. Our entire existence amounts to a long and ceaseless compromise between what we could do and what is within our reach, between what is conceivable and what is actually possible. And the exact nature of this compromise depends on how strictly we choose to define personal freedom. If the latter is construed in a tolerant manner, as the relative power to determine our choices and actions, there may still be considerable room for autonomy and independence. But if freedom is defined, *sensu strictiore*, as the absence of each and every kind of limitation, our lack of independence will appear even greater. For even our most modest wishes and projects will be found to depend on an endless series of constraints, related to the shape and strength of our body, to the physiology of our brain, to the spatial and temporal structure of the surrounding world, etc.

For every human being, freedom is confined to an interval located somewhere between complete powerlessness and omnipotence. The interval’s upper end determines the actual extent of personal freedom, but its position is usually open to variation. Although most choices are indeed very limited, not all limitations are equally constraining: while some are more easily endured, and sometimes even hardly noticed, others represent more serious threats to the daily running of human life. In light of this variation, the different obstacles to personal freedom can be compared, in Stoic vein, to the walls of a citadel. At a more superficial level, the acknowledgement that one is not free to do anything one chooses is not particularly

²Ibid. The poem highlights three ‘words of faith’, namely Freedom, Virtue and God (or a divine Will), and Schiller argues that ‘Mankind is of all [its] value bereft / If in these three words no faith is left.’ Although I am presently concerned with the first of these words, the following sections will show that the issue of human freedom is much wider than usually supposed: when human beings ‘lose faith’ in freedom, they are also at a loss with respect to virtue (or ethics in general) and God (or transcendence in general).

hard to accept and still leaves room for a vast array of possibilities. Although human beings are unable to fly, for instance, or travel through time, they have come to cope with these limitations and live their lives accordingly. Further inside the citadel, however, freedom's interval can grow much smaller. If men and women are prevented from speaking freely, or sent to jail, or sold into slavery, the balance between what they can and cannot do is significantly altered. Worse still, if they are denied the control of their body, by being gagged or shackled, their freedom is compromised even further, with direct implications in almost every aspect of their life.

Yet these constraints, however serious, are never serious enough to invade one's inner citadel and destroy one's implicit belief in personal freedom. Which brings us back to Schiller's claim: although the actual power to choose and act freely may be wider or more restricted, depending on the number and nature of the obstacles one is up against, it nonetheless *presupposes* a minimum of self-determination. Slaves or prisoners, despite all their limitations, can still cling to a more basic form of autonomy. They can still regard their bondage as an accidental predicate, added to a more essential sphere of personal freedom.

This basic sense of autonomy is the ground zero of human freedom and the foothold of every human choice. Whereas all other forms of freedom, once denied, still leave room for some degree of independence, this one does not. It is the element that holds everything together, the inner core that grants meaning to freedom's outer layers. Therefore, if it should prove unreliable, the entire interval of freedom would be compromised. If we did not think of ourselves as fundamentally free agents, we would not experience the world as we normally do: the very notion of possibility would be transfigured and life would no longer appear to us as an open process, or a task awaiting completion. In these circumstances, the ability to choose one thing over another would no longer have any meaning because it would no longer belong to *us*, but to *something* or *someone else*: even if all other impositions were suddenly lifted, and we were allowed to do whatever we please, this new-found power would still amount to a deception.

In its most basic form, then, our belief in personal freedom amounts to the feeling that we *own* our life, that we are the authors of our thoughts and actions and not mere subjects of an alien power. And this belief is reiterated at each waking hour. Time and habit strengthen the conviction that our whole existence, despite its considerable share of contingency and impotence, is steered by us; that we are its protagonists, as it were, and that its varying circumstances, undoubtedly real and effectively constraining, are ultimately but an outward setting, which can either favour or obstruct our natural sovereignty.

But this is not all. While freedom is usually measured by what we can and cannot do, it is not limited to this alternative. Apart from the power to choose and act according to our will, freedom also requires a certain degree of insight into the actual import of our choices. Indeed, freedom is just as much about power as it is about knowledge: for our choices to be free, they must be meaningful, intelligible and coherent choices, and for that to be so we must be able to make sense of reality and interpret, as best we can, what goes on around us.

A basic characterization of personal freedom must not be limited, therefore, to its concrete requirements, related to the kinds of things we can and cannot do. Such a characterization must also take into account freedom's *cognitive* requirements, and the kinds of things we know and ignore. Once more, these requirements can be compared to different layers, or different walls, built around an inner core of cognitive self-confidence. Like personal power, personal knowledge is usually confined to an interval placed somewhere between complete ignorance and omniscience. The interval's upper end determines how knowledgeable we actually are, whilst the lower end indicates the minimal insight required for the exercise of freedom. Just like before, not all forms of ignorance are equally essential or equally constraining: our ignorance of quantum physics, for example, is liable to affect our power to make informed scientific decisions, but the inability to read or write can prove far more damaging, as it is bound to condition our freedom of choice in a wider number of situations; and if we ignore, moreover, the most basic principles of the physical world, we will be even less able to determine and direct the course of our lives.

Moving further down the interval, we arrive at a new ground zero, whose suppression would entail the suppression of freedom itself. Once again, although our independence may be wider or more restricted, depending on the number and the nature of the things we know, it presupposes a minimum of knowledge. At the very least, free choices must be grounded in a general assessment of what is *at stake*, at each given moment, in our lives. And though this assessment may be vague, volatile or even unconscious, it is the foundation of a whole edifice of cognitive assumptions, whose inner composition is no less important to the definition of freedom. In particular, if we have no idea of who we are, where we stand or what is happening around us, our decisions can hardly be called free. For something to be freely decided, there has to be some kind of awareness of the subject, object and context in which the decision is made. Furthermore, apart from this basic knowledge, free choices also require the power to determine, in each situation, which and how many different alternatives are available to us. Again, if this power is lacking there can be no real freedom, for our choices are bound to reflect our ignorance rather than our actual intentions. Finally, personal freedom is also conditional on the ability to anticipate, at least up to a point, the consequences of our decisions. We may have a clear picture of the situation we are in and of the best possible course of action and yet be surprised by the actual result of our choices. In this case, although we are free to choose, we may end up getting the opposite of what we bargained for.

To appreciate the extent to which freedom is dependent on these basic requirements, and the extent to which they are dependent on each other, let us consider what would happen if each of them were removed. If human beings, when faced with a given choice, were able to identify the different alternatives available to them and to anticipate their consequences but utterly unaware of their own identity, the first two powers would be worthless, for the freedom they afford would be grounded in ignorance. This kind of freedom would be as contradictory as that of a blind man

allowed to see what he pleases. Alternatively, if human beings were able to make out their practical situation and to anticipate the effects of their choices but unable to determine which and how many options were open to them, the first two powers would also be compromised. This combination would result in a new form of blindness, whether or not it were perceived as such. Thirdly, if we knew who we are, where we stand and what options we have before us but were utterly unable to foresee where they lead, our freedom would also be illusory, for we might spend our whole life choosing and acting against our interests. In this case, our vision would amount, yet again, to a peculiar form of blindness.

Since all of these clauses are essential to the actualization of freedom, the removal of one, two or all three of them would result in impotence and blindness. Such an exercise is typically enacted in Greek tragedies, where personal freedom is often countered by ignorance and self-deception. Oedipus' case is the most exemplary: although most of his decisions are freely made, they rest upon a globally mistaken diagnosis of the situation he is in; and though he is convinced he has a relatively clear knowledge of the different elements at play in his drama (viz. his own identity, the identity of his parents, the identity of his wife), he soon finds out that nothing is what it seems (viz. that his adoptive parents are not his biological parents, that his mother is also his wife, that he himself is the killer of his father). Sophocles' tale is exemplary because of its unique focus on freedom's cognitive dimension. Oedipus embodies the extreme possibility of a man whose actual freedom to choose and act is completely neutralized by his cognitive incompetence. Although he is able to do what he pleases, his choices are accompanied by a complete lack of insight into his actual situation, a complete ignorance regarding the options that are actually open to him and a complete inability to anticipate their consequences.

The main point, however, is not simply that Oedipus knows not what he thinks he knows, but also that his power, due to his ignorance, is not as real as he takes it to be. In the play, this idea is conveyed by Sophocles' ingenious interweaving of freedom's concrete and epistemic claims. Whereas in the first scenes Oedipus' freedom appears to outshine everyone else's, his superiority is soon called into question. As the play unfolds and his ignorance becomes more apparent, his ability to shape circumstances in his favour is increasingly challenged. Oedipus' choices are grounded in a flawed diagnosis of his personal situation, which leads to an equally flawed prognosis of its future developments. And although his troubles are initially put down to outside forces or unfavourable circumstances, they are ultimately due to an overestimation of his own cognitive powers.

This whole idea can be translated in Aristotelian terms, by the tragic notions of recognition (ἀναγνώρισις) and reversal (περιπέτεια).³ Sophocles' play is designed in such a way as to maximize the impact of both of these elements. As regards the first one, Oedipus' tragedy does not consist merely in having to acknowledge that he is deluded about this or that particular aspect of his life. What is in store for him is not a partial ἀναγνώρισις, but a *total* one. And this is all the more remarkable since

³See Aristotle, *Poetics* 1452a12ff.

Oedipus has already been informed by the oracle of what is going happen. He already knows the truth but still does not see it, and his extreme situation is meant to illustrate just how complete human ignorance can be.⁴ As regards the second element, Oedipus' peculiar situation does not lead to a simple change of fortune, but to a reversal of his entire fate. Again, what is at stake at the end of the play is not a partial *περιπέτεια*, but a *total* one, whose seriousness is meant to illustrate just how flawed human reason can be. Out of ignorance, Oedipus is led not only to misidentify all the elements at play in his drama, but also to misrepresent the practical implications of all his decisions. In the absence of a suitable diagnosis of reality, his alleged power veers round into impotence and servitude.

Yet despite Sophocles' eloquence, it may be argued that Oedipus' tale is overly dramatic or unrealistic, that Sophocles exaggerates the importance of freedom's cognitive dimension, or that his convoluted plot has no equivalent in real life. Just as Oedipus reacts with incredulity and scepticism to Tiresias' predictions, so too most of us tend to reject the idea that our knowledge of reality amounts to an illusion, or that our freedom can be overturned in such a complete and catastrophic way. Whereas Oedipus' whole existence is based on a series of mistakes, our lives are thought to be less open to deception. And whereas Oedipus' freedom is made to depend entirely on what he knows or ignores, our own freedom is thought to hinge on more concrete circumstances. Interpretation and anticipation are generally regarded as relevant conditions for the actualization of freedom, but their importance is only fully acknowledged in connection with complex dilemmas or difficult choices. And while the latter are certainly real and highly significant, they tend to represent but small islands in the wider sea of our daily concerns.

This reaction, however plausible, is really only a confirmation of the anatomical picture I have been sketching. It is precisely because none of us is ever indifferent to the issues of whether and to what extent we are free that Sophocles' tale seems so far-fetched. In our own case, we tend to assume we are knowledgeable enough to lead relatively free lives and to shape most situations according to our interests. Yet Oedipus is just as attached to his own cognitive perspective and just as convinced of his freedom. In this regard, Sophocles' portrait is indeed an accurate illustration of human nature. For all of us, as for Oedipus, true freedom is the opposite of tragic freedom: it requires a kind of knowledge that cannot be reversed. But just as Oedipus was eventually disabused of his cognitive self-confidence, so too may we be exposed to a similar fate.

⁴The tragedy (or the irony) does not lie simply in the fact that Oedipus ignores what he thinks he knows. The very provisions he and the other characters make to avoid the fate prophesied by the oracle—namely, that Oedipus will kill his father and sleep with his mother—end up actively contributing to his downfall. For instance, Oedipus' departure from Thebes to avoid fulfilling the oracle's prophecy leads him to his home town and to his real parents; Jocasta's surrender of her young son to escape the same prophecy sets off the chain of events that will lead to murder, incest and suicide. In these situations, the characters' intentions are not merely opposed to the actual outcome of their choices: they are the vehicle whereby that outcome is brought to pass.

The former description suggests a few important conclusions. To begin with, we have seen that personal freedom is not a simple, but a complex phenomenon. The usual assumption that we are free—or at least as free as we need to be in order to do what we happen to be doing—rests on two important sets of presuppositions, which correspond to different forms of power. However, their differences are usually mingled into one blurry image. When I claim, for instance, that I am free to leave this room, what I am saying is that ‘I have the power, if I so choose, to leave this room.’ And this requires in turn (a) that I am the primary source of my thoughts and actions, and not a puppet steered by an unknown master; (b) that I can decide whether to stay or to leave this room; and (c) that I am capable of carrying out my decision. Moreover, since personal freedom is also dependent on a series of cognitive presuppositions, the power to leave this room also requires (d) that I am aware of what is at stake in my own life; (e) that I have a clear enough knowledge of my present circumstances; and (f) that I can anticipate the consequences of my actually leaving this room.⁵

Furthermore, although all of the powers usually entailed in the experience of personal freedom are equally important, they are not equally fundamental. Apart from the inner anatomy of our belief in personal freedom, we must also consider its inner physiology. The different powers we take for granted are not merely coordinated with each other, but also subordinated to one another, with the most elaborate presupposing the most simple. Returning to the previous example, I am only free to leave this room because I have the power to anticipate the consequences of my choosing to do so. But this power is conditional on my ability to decide whether to stay or to leave, which in turn requires a minimal understanding of the practical situation I am faced with. Finally, none of this has any meaning if I am a mere pawn of fate or the instrument of an alien will. My belief in personal freedom requires, on the contrary, that I be the starting point of my choices and that I recognize myself as such.

All of these powers are at work, on different levels, in my usual conviction that I am free, if I so choose, to leave this room. And they are also present in most of the choices and decisions made throughout my life. However, not all of them are as certain or self-evident as they appear—which brings us to the third and most important point of our enquiry. The ordinary conception of personal freedom rests on a series of beliefs automatically held to be true. However, since their truthfulness is usually merely taken for granted, and not actually proven, it is open to various kinds of criticism. Considering what went before, this means not only that each of the powers we take for granted may turn out to be illusory, but also that *all* of them, inasmuch as they presuppose one another, may turn out to be illusory.

⁵ At first sight, the awareness of what matters most in life is not needed to decide something as trivial as leaving this room. The determination of one’s life goals is a complex and profound issue, whereas the decision to leave a room is much more straightforward. But though we usually fail to notice it, the decision to leave this room—like any other human decision—is necessarily supported by a wider, if inexplicit appreciation of our whole life. This decision is never meaningful or desirable in itself, but always as part of a wider existential script, as it were, grounded in a million other decisions, assumptions and expectations. I will return to this issue in Sect. 2.2.

Let us start with freedom's concrete claims. As regards the most basic assumption in my previous inventory, there are compelling philosophical arguments against the idea that we are the original source of our thoughts and actions. According to some brands of determinism, the feeling of self-control that runs like a *basso continuo* through our entire lives is not a sufficient proof of our independence. According to other, more radical brands of determinism, this feeling is downright illusory and we are indeed mere links in a predetermined causal chain.

Furthermore, concerning the actual power to choose a specific course of action, we may reject the previous modes of determinism and still find that our independence is extraordinarily limited. Even if we are not an intermediate link in a predetermined causal chain, we may be the initial link of an equally inexorable chain, in which case our particular choices are not free choices, but mere reactions to previous events. Although we appear to be in control of our minds and bodies when deciding to move around, or speak, or pour a cup of coffee, we may simply be responding to highly complex arrangements of physical, biological and social stimuli. And if it were possible to isolate all of them, to determine their interconnection and to account for all their effects, we might very well find that our choices are not truly voluntary.

Finally, to complicate matters even further, our belief in the ability to carry out our decisions and achieve our projected goals is no less problematic. Even if we are indeed the original source of our thoughts and actions, and capable of choosing freely and acting upon our choices, there is ultimately no way of guaranteeing their success. For though we may be free to set in motion a determinate sequence of events, its final outcome is not exclusively up to us. It is dependent on a constellation of factors that cannot be fully anticipated or controlled, including the laws that govern the natural world, the choices made by other people or the dictates of chance.⁶

So much for freedom's concrete dimension. Let us now return to its cognitive dimension. As seen earlier, although we are often faced with situations we do not understand, our grasp of the world around us is usually deemed good enough to inform most of our decisions. And though we often recognize we are wrong or misinformed, our mistakes seem to call for small, localized corrections, and not for a global revision of our entire outlook on reality. But how solid are these assumptions? Are they open to refutation? And if so, to what extent?

The purpose of this book is to help answer these questions. So far, I have argued not only that our usual belief in personal freedom is complex and multi-layered, but also that most of its underlying presuppositions can be called into question. In what

⁶When shooting an arrow, for example, the probability of hitting the target can be increased by practising our aim, calculating how far back the bow's string should be stretched, determining the speed and direction of the wind, etc. But despite all of these precautions, we will never be able control all the variables that might affect the success of our task. An earthquake might suddenly occur, a momentary cramp might limit our movements, etc. Due to these and a million other possibilities, the efficaciousness of our choices is always an open issue, which means that the real extent of our freedom is much narrower than we are usually prepared to assume.

follows, I will not take up the whole issue of personal freedom. I mean to focus, in particular, on the relationship between personal freedom and personal knowledge. More specifically, I mean to question the assumption that we are knowledgeable enough to lead free lives, and to do so I will start by focusing on the nature, structure and scope of our usual understanding of reality. Revisiting the history of Western philosophy, I will argue that our cognitive sovereignty is not as certain as it appears, nor is the freedom it is thought to provide. Moreover, I will show that this is not merely a theoretical problem, but one that has direct implications in every aspect of human life. Finally, I will argue that true freedom requires a radical revision of our usual standpoint and the pursuit of a new cognitive perspective.

1.2 Freedom and Education

The previous considerations have led to a conclusion that will remain at the centre of this enquiry: namely, there is no freedom without cognitive awareness. To be sure, our usual insight into the meaning and consequences of our choices is never fully adequate. No matter how perceptive we take ourselves to be, our cognitive powers can hardly match the extraordinary complexity of human life. However, if we had no idea of who we are, where we stand or what goes on around us, our choices would have no meaning. Living would be like walking through a darkened room, with no familiar references to guide the way. Therefore, we usually take a middle course: although we do not expect our knowledge to be perfect, we do expect it to provide us with a sufficient degree of independence. We require a knowledge powerful enough to light up the darkened room of our lives and reveal, if only faintly and within a small perimeter, the shape and colour of the surrounding world, the direction in which we are heading and the main obstacles along the way.

The debate concerning the nature and the amount of knowledge required for the actualization of freedom is an old and difficult one. Historically, its origin can be traced back to the ancient Greek contrast between human beings' primitive or natural standpoint and a civilized or educated standpoint. The Greeks were the first to insist on the idea that freedom is not an immediate ontological asset, but the result of a specific form of education, or *παιδεία*, destined to release human beings from their natural state of ignorance. As the word suggests, *παιδεία* was the transitional process whereby the *παῖς*, i.e. the child or youth, was elevated to the state of *πολίτης*, i.e. citizen, and admitted into the city's public and political life. But the word *παῖς* was also used for slaves and servants in general.⁷ It referred to a kind of immaturity that was not exclusively tied to one's age, but also, and primarily, to one's power of judgement. Children and slaves—but also women and mad people—were all

⁷An ambiguity still found in the English word 'boy' or the French word 'garçon'. See Aeschylus, *Choephoroi*, l. 653; Aristophanes, *Acharnians*, l. 395, *Clouds*, l. 132; or Demosthenes, *Against Apaturius*, 33, l. 8.

included in this category because they were all deemed intellectually immature, and hence unaccountable for their judgements and actions. Lacking the emancipation provided by παιδεία, they were denied the right to free citizenship and destined to a life of submission.

The history of Greek education is based on the antithesis between cognitive maturity and cognitive immaturity. From the more primitive to the more sophisticated versions of παιδεία, what distinguished the educated (παιδευτός) from the uneducated (ἀπαιδευτός) was the acquisition of the knowledge deemed appropriate to a free or independent life—the knowledge required to make one's own decisions, to shape circumstances in one's favour and to take part in the life of the community.⁸ With time, this formal distinction was adapted to different contents and subordinated to different cognitive requirements. The kind of insights that were initially deemed sufficient to ensure one's independence were eventually found wanting, which prompted the need for more sophisticated forms of education. As human life grew more complex, the emergence of new needs, habits and aspirations led to the redefinition of the frontier separating education and freedom from ignorance and servitude.

The first great step away from humankind's natural ignorance is described by many Greek authors in mythological terms, as the result of a divine gift. As Prometheus claims in Aeschylus' (or Pseudo-Aeschylus') homonymous play, although primitive men 'had eyes to see, they saw to no avail; they had ears, but they did not understand; . . . just as shapes in dreams, throughout their length of days, without purpose they wrought all things in confusion'.⁹ Only when they were handed down the gift of fire and the secret to a series of practical arts (τέχναι) or skills (σοφίσματα)—such as building, writing, fishing or mining—were they freed from their original blindness and 'endowed with reason'.¹⁰ This same myth is evoked by Protagoras in Plato's homonymous dialogue, where Prometheus is also credited with having imparted to the human race the knowledge of the various arts.¹¹ When the gods created the living world, they provided all animals with speed, or strength, or a natural habitat, or other means of survival and preservation. Human beings, however, were left 'naked, unshod, unbedded, unarmed'.¹² To compensate for this oversight, Prometheus offered them the gift of fire and the secret of all of Athena's arts, thereby ensuring their security and livelihood. A brief catalogue of these endowments, similar to the one offered by Aeschylus, can also be found in the first stasimon of Sophocles' *Antigone*, where the power of human reason is said to

⁸Part of this approach is still valid today. In most modern societies, the intellectual immaturity of children and the cognitive limitations of mentally ill and impaired individuals are usually regarded as a challenge to their natural freedom and the basis for certain legal restrictions. However, in contrast to the Greek model, those individuals are nonetheless fully fledged citizens, endowed with a full set of basic human rights.

⁹Aeschylus, *Prometheus Bound*, ll. 447–450

¹⁰*Ibid.*, 491

¹¹*Pr* 320cf.

¹²*Pr* 321c5–6

‘span the seas’ and ‘wear away the earth’, to ‘turn the soil’ and ‘snare birds and fish in the meshes of twisted nets’, to ‘master the beast who dwells in the wilds’ and ‘flee the arrows of the inhospitable frost’.¹³

All of these accounts accentuate the extraordinary transformation brought about by the advent of reason and education. Whereas primitive men and women were ignorant and defenceless creatures, vulnerable to all sorts of constraints, educated men and women are stronger, cleverer, more industrious and more organized. They can challenge nature’s tyrannical rule, anticipate and influence the course of natural events and harness the various resources of the physical world. At the same time, however, their new understanding of reality gives rise to more sophisticated cognitive demands. As their knowledge increases, so does the awareness of their limitations and the need for new forms of instruction.

In the ancient Greek world, this revolution was accompanied by another, closely related one, brought about by the so-called sophistic movement. With the sophists, education acquired a new importance and a broader scope. Building on the Promethean leap from ignorance to knowledge, the sophists highlighted the possibility of a new and equally radical leap forward, towards an even higher cognitive stage. Just like before, however, the result of this new transition can only be truly grasped at the point of arrival. This movement can be interpreted in light of what the German classicist Hermann Fränkel called a ‘Heraclitean thought pattern’, that is, the analogical procedure whereby the contrast between a lower term X and a higher term Y is used as reference to envisage the identity of an even higher term, Z, which relates to Y in the same way that the latter relates to X.¹⁴ This kind of analogy is usually employed as a means of anticipating the content of ideas or realities that lie beyond one’s cognitive reach: in this case, starting from the familiar contrast between ignorance and the basic resources available to most educated individuals (X and Y, respectively), one is led to envisage the unknown contrast between this basic set of resources and a new kind of knowledge (Y and Z, respectively), the attainment of which would amount to an equally profound cognitive upgrade.

In other words, the sophists highlighted, with unprecedented vigour, the intrinsic *mobility* of human cognition. They aimed to show that human beings are not confined to a single cognitive perspective, or to a single outlook on reality, but capable of moving beyond their native standpoint and embracing new and previously inaccessible forms of knowledge. Παιδεία, in the sophistic sense, stands above all for a cognitive transfiguration: whereas education was previously construed in a chiefly quantitative manner, as the possibility of expanding a basic repertoire of skills or resources, the sophists emphasized the qualitative gap separating this basic standpoint from a truly educated standpoint. Challenging the immediate accessibility of the knowledge handed over by the gods, they replaced the general distinction between skilful and unskilful individuals with the stricter distinction between experts and laymen.

¹³ Sophocles, *Antigone*, ll. 332ff.

¹⁴ See Fränkel, ‘A Thought Pattern in Heraclitus’, 314ff.

This attitude is the starting point of a cognitive trend whose effects are still recognizable today. Although modern men and women tend to rely, in most daily situations, on the knowledge provided by experience and common sense, they are well aware of its limitations. What most people know about a certain subject or phenomenon is usually assumed to be different from what an expert knows about it: his or her insight is not just any insight, but a specialized one, which exceeds human reason's basic cognitive repertoire. While most of us know, for example, the general difference between a headache and a toothache, not all of us are familiar with their exact causes or with what it takes to stop them. And although most of us know the difference between a house and a bridge, few of us know how to design either of them, or how to build them and make them stand. These insights are usually the prerogative of doctors, dentists, architects and engineers. They amount to artificial extensions of human reason, exclusively available to those who have studied and exercised themselves, often for several years, in these activities.

In the sophist's vocabulary, doctors and architects who know their trades are in possession of a *τέχνη*, that is, an art or expertise enabling them to know what most people can only guess. Experienced doctors do not stare powerlessly at the suffering of their patients because they are able to see *beyond* its immediate appearance. While disease and decay are for most of us, non-doctors, a mysterious affair, whose exact origin lies somewhere inside our body, a doctor's interpretation of the same phenomena is very different. Where we see a regular human body, a doctor sees a highly complex system of muscles, bones, sinews and vessels. And where we stare alarmed at the swelling of the skin or the onset of mysterious convulsions, a seasoned doctor instantly recognizes the usual symptoms of a lipoma, or the unmistakable signs of neuronal hyperexcitability.

Doctors, dentists, engineers and other *τεχνῖται* see the world differently. What is more, they can act upon it in different ways. By determining the causes of certain phenomena, they are free to manipulate them and to change their effects. Conversely, those who lack this kind of insight are less able to shape circumstances in their favour. They are doomed either to suffer the consequences of their ignorance or to subject themselves to the authority of others. This sophistic theme is evoked, for example, in Plato's *Lysis*, when Socrates claims that 'with regard to matters in which we become intelligent . . . we shall do as we please, and nobody will care to obstruct us. Nay, not only shall we ourselves be free and have control of others in these affairs, but they will also belong to us, since we shall derive advantage from them; whereas in all those for which we have failed to acquire intelligence, . . . we on our part shall be subject to others.'¹⁵

On the political level, the sophistic parallel between knowledge and power led to an increase in the social importance of education, which prompted a redefinition of the rights and duties associated with free citizenship. In his classic study on ancient

¹⁵ Ly 210a9–c4

Greek education, Werner Jaeger points out that the epithet ‘free’ (ἐλεύθερος), in the fifth and fourth centuries BC, usually meant that one is ‘not a slave’ (δοῦλος, παῖς), while ‘the kindred word “liberal” (ἐλευθέριος) describes the conduct appropriate to a free citizen, whether in generous spending or in frank speaking (which would be improper in a slave), or in a gentlemanly way of life.’ Accordingly, ‘the “liberal” arts are those which belong to “liberal” education—and that is the paideia of the free citizen, as opposed to the uncultured vulgarity of the unfree, of the slave.’¹⁶

Following the sophistic model, the notion of παιδεία came to stand for an increasingly wide array of practical and theoretical subjects, ranging from gymnastics and military training to grammar, rhetoric, music and medicine. This curriculum reflected both an individual and a social commitment to freedom: individually, it enabled young men to overcome their natural weaknesses, by helping them improve their physical and cognitive powers; socially, it aimed to guarantee the security and welfare of the polis. Upon reaching adulthood, educated individuals were accorded the title of citizens and encouraged to become statesmen, to embrace a military career or to pursue a liberal profession. In all of these activities, they were expected to put their expertise at the service of the city and to help defend it against natural and political threats, to increase its economic and military might and to enrich its cultural and intellectual life.

By putting a new premium on knowledge and education, the sophists raised the frontier separating cognitive minority from cognitive majority and placed higher demands on citizens and free men. They maintained, not without controversy, that the πολῖται were still, in a certain sense, παῖδες—that they were not as free or insightful as they believed, but still in need of education. Yet although the sophistic revolution amounts to a radicalization of the primitive notion of παιδεία, its scope is still very restricted. Although the sophists highlighted the limitations of ordinary knowledge and the possibility of more sophisticated cognitive perspectives, their critical efforts were themselves targeted by a new and more extreme educational project, outlined and discussed in Plato’s dialogues.

Plato’s idea of education is also linked to a radical cognitive reform, but its scope is very different from that of the previous models. Like the sophists, Plato is very critical of human beings’ usual cognitive perspective, but his criticism does not apply only to primitive individuals and educated Greek citizens. It also extends to the sophists themselves, as well as to experts of all kinds. To be sure, he does not regard all human knowledge as equally valid, nor does he believe every human being to be equally knowledgeable or cultivated. Yet despite these variations, Plato highlights the possibility of a more general and more serious form of ignorance, shared by every man and woman—a form of ignorance that lies at the very heart of human cognition, affecting both the general knowledge of laymen and the specialized knowledge of doctors, engineers and other τεχνῖται.

¹⁶ Jaeger, *Paideia*, vol. 2, 54f.

The debates between Socrates and his interlocutors are centred, in most cases, on the definition of a familiar word or concept, whose meaning is usually taken for granted.¹⁷ As the dialogues unfold, this assumption is shown to be groundless and the first attempt at a definition gives way to a second, and to a third one, and often to many others. All the while, Socrates' task is simply to question his opponents, compare their answers with one another and point out their contradictions. At first, there is no reason to suspect that human cognition is fundamentally flawed. If the debaters are unable to come up with a valid definition of a given concept, that only means that the latter is more complex than initially supposed. To be properly defined, it must be broken down into more basic components. The exercise is thus repeated at a more elementary level: new concepts are brought up and new definitions are called for, but the results are equally discouraging. If the initial debate revolves around the notions of justice and injustice, for example, or beauty and ugliness, the interlocutors soon find themselves grappling with much simpler determinations, such as 'good' and 'bad', 'large' and 'small', 'fast' and 'slow'. And here, again, what seemed familiar and straightforward turns out to be neither. Eventually, they are led to focus on notions as elementary as unity, identity or difference. But even here a consensus is hard to obtain. Just like before, these notions prove ambiguous and ultimately unintelligible.

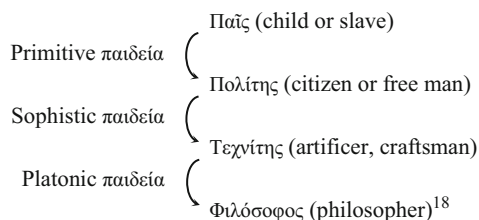
What emerges from this exercise is not merely the problematic nature of this or that human concept or set of concepts. Plato is making a much more extreme suggestion, namely that human reason is *globally* problematic. He is suggesting that the validity of every human belief is dependent on a series of concepts which are themselves merely assumed to be valid; and that the latter are supported, in turn, by a further subset of similar assumptions. Worse still, he is suggesting that the world's usual familiarity is grounded in a set of concepts which are themselves unfamiliar, and ultimately unintelligible; and that the latter are grounded in a further subset of similar assumptions. This extreme diagnosis amounts to a radicalized version of Sophocles' hypothesis: like Oedipus' knowledge, human knowledge is grounded in a web of assumptions that lend force to each other, but whose actual truth is far from certain. And as in the play, to question the validity of any one of these assumptions is to unveil a much wider cognitive problem, whose seriousness increases with the depth of one's enquiries.

Plato's critique of human cognition will be discussed in the following chapters. For now, my purpose is to highlight the specificity of his conception of *paideia*. In the previous stages, the transition from a less educated to a more educated standpoint involved a moment of self-criticism. The most primitive form of education followed the acknowledgement of mankind's original impotence and the need for new and more sophisticated cognitive resources. Later, the sophists highlighted new forms of impotence and the need for more sophisticated accounts of reality, irreducible to the

¹⁷ Throughout this work, and unless otherwise noted, all references to Socrates concern not the historical figure but the character portrayed in Plato's dialogues, viewed as the central mouthpiece of Plato's philosophical project.

general dictates of experience and common sense. Plato's educational reform can also be included in this general movement. Like the previous moments, it is also grounded in a critique of human cognition, but one that encompasses all of the previous educational achievements. Moving beyond the practical teachings of Prometheus and the 'technical' teachings of the sophists, Plato highlights the need for a *philosophical* reform of human cognition—the only one capable, in his eyes, of exposing and reversing human beings' cognitive limitations.

This entire progression can be summed up in the following scheme¹⁸:



Plato's philosophical project can be regarded as a new chapter in the history of education, and it differs from the previous ones in at least two fundamental respects. The first one, already mentioned, is the wideness of its scope. At stake in Plato's enquiries is a truly universal form of education, directed not at a specific kind of individual or group, but at *every human being as such*. And this aspect is itself of great historical and cultural significance, as it introduces a cosmopolitan approach to human nature that is relatively new. In ancient Greek culture, human relations were not yet grounded in the modern biological notions of 'humanity' or the 'human species'. Ancient Greek men recognized others as their equals on the basis of their birthplace, language or social rank and drew very clear distinctions between citizens and non-citizens, compatriots and foreigners, men and women, etc.¹⁹ In Plato's dialogues, however, human ignorance is presented as an essential human trait, which transcends racial and political distinctions. Although the latter are still at work in different and important ways, they are overshadowed by the *universal fact* of human ignorance. Education, rather than politics or biology, is the true universal denominator, binding every man and every woman to the same existential predicament.

Furthermore, the wideness of Plato's critique does not refer only to its target, but also to its actual object. As we have already started seeing, it extends to the most basic elements of human cognition, and hence to every concept, idea or judgement relied upon in everyday life. Indeed, Plato's aim is not merely to identify and correct

¹⁸This scheme is of course a simplification. The different forms of παιδεία amount to fluid, interrelated cultural events, which cannot be reduced to a simple definition or a precise historical period. Not only was the first form παιδεία influenced, early on, by the sophistic model, but the latter was also instrumental to the development of Plato's philosophical model.

¹⁹See notably *St* 262c2ff.

certain forms of ignorance, or to secure the knowledge required to achieve certain ends. In his dialogues, human cognition emerges as a globally aporetic phenomenon, in need of a global rehabilitation.

The second main novelty of Plato's approach concerns the method employed in his dialogues. In the *Sophist*, for instance, when discussing the issue of education, Socrates explicitly evokes the Athenian conception of παιδεία, but only to reinterpret its conventional meaning: instead of defining education in a positive manner, as the acquisition of an increasingly specialized understanding of reality, Socrates defines it in a negative manner, as the effort to 'cast out the conceit of knowledge'²⁰ underlying most human thoughts. He argues that the most grievous form of ignorance, 'separate from the rest, and as weighty as all other parts put together', is 'thinking that one knows a thing when one does not know it'. 'Through this', he adds, 'all the mistakes of the mind are caused in all of us.'²¹

Plato's critique is a *negative* critique: instead of expanding human knowledge, it aims to contract it, or rather to roll back the pretence of knowledge that pervades human life. Ignorance is thereby exposed as a widespread and involuntary phenomenon, whose first 'victims' are precisely those who judge themselves more knowledgeable. This unflattering conclusion, reiterated throughout the dialogues, calls into question the validity of every traditional form of education, whether based on practical, political or technical standards. It suggests that every human being, however experienced or powerful, is still a mere παῖς, viz. a mere infant, or a mere slave.²²

²⁰ So 230b1. The word δόξα can be translated as opinion, belief, judgement, etc. In Plato's works, however, it is usually associated with an *unwarranted* opinion, belief, or judgement. I will return to this issue below.

²¹ So 229c1–6

²² It might be argued that this is true of Socrates' educational method, with its focus on irony, ignorance and the disavowal of knowledge, but not so much of Plato's, which appears to be more positive. After all, the *Republic* does elaborate on the curriculum to be taught to future philosophers and political leaders. In what follows I challenge this distinction by showing that Plato's conception of paideia, whether in its so-called 'Socratic' or properly 'Platonic' versions, is *constitutively* grounded in a global critique of human cognition. Drawing on the *Republic* and other texts, I argue that Plato's use of dialectics, even when seemingly least polemical, is never simply about the acquisition of knowledge, but always about the refutation of an existing pretence of knowledge—not about 'putting sight into the eyes of the blind', but 'turning the soul around' and 'devising a way of making the organ see, which already has vision, but is not . . . looking where it should.' (*Re* 518c1–d7).

This issue leads to the wider debate on the differences between Plato's early, 'Socratic' dialogues, explicitly aporetic, and his later, 'Platonic' ones, of a more conclusive nature. Again, while I do not deny the heterogeneity of the dialogues, I will show that they are never as conclusive as they are made to appear. Moreover, I will argue that the tendency to find definitive positions and solutions in the dialogues is usually coupled with the tendency to portray Plato as a dogmatic metaphysician, committed to a clear separation between the sensible world and a divine or heavenly world of ideas. One of the aims of this book is precisely to counter this tendency (see Sect. 1.5). In my view, only by moving beyond this rigid dualism and acknowledging the *immanent* tension highlighted in the dialogues can one appreciate the true implications of Plato's philosophical criticism.

The most famous illustration of this idea can be found in the *Republic*'s seventh book, where human life, 'as regards both education and the lack of it',²³ is compared to a peculiar form of captivity. Socrates imagines a group of prisoners sitting in the bottom of an underground cave, forced to face forward and to watch the shadows of different objects projected on a wall. The prisoners' legs and necks are bound by shackles, which prevents them from seeing each other, their own bodies, or indeed anything other than the shadows on the wall. Moreover, since their imprisonment has lasted for as long as they can remember, their lack of freedom is not perceived or experienced as such. They do not know any other kind of life, and live therefore under the assumption that they are free and that the wall in front of them encompasses the whole world.

According to Socrates, these prisoners are 'just like us',²⁴: their ignorance resembles our own ignorance and their captivity illustrates the lack of freedom imposed on us by our usual understanding of reality. In other words, just as the prisoners are ignorant and unfree but fail to notice it, we too are ignorant and unfree but fail to notice it; and just as the prisoners, in order to acknowledge their captivity, must break their chains and look around, we too must free ourselves from our usual pretence of knowledge. Once again, this image is grounded in a Heraclitean thought pattern, designed to convey a possibility that is usually out of reach: starting from the familiar contrast between the prisoners' life and our own life, we are led to envisage the unknown contrast between our life and a new kind of life, whose embracement would amount to an equally profound cognitive transformation.

As Schopenhauer wrote, 'every man regards the limits of his own vision as the limits of the world.'²⁵ The allegory of the cave is an eloquent illustration of this. Since the cavemen are unable to see the chains that bind them, they mistake the cave for the whole world. Therefore, their education seems to be, for the most part, a closed affair: although their learning is never really over, for there are always new shadows to see and new meanings to interpret, their overall conception of reality is not likely to be radically transformed. In like manner, inasmuch as most human beings believe their usual understanding of reality to be fundamentally reliable, they do not bother to question its underlying principles. Indeed, although human reason is far from perfect, it is usually deemed accurate and empowering enough to allow for a relatively free life. Although we never cease to learn new things, or to correct and refine our previous knowledge, we do not expect our usual standpoint to be radically refuted or transformed. According to Plato, however, this expectation may be entirely unwarranted.

²³ *Re* 514a2

²⁴ *Re* 515a5

²⁵ Schopenhauer, *Werke*, vol. 5, 245

1.3 Plato, Hegel and Nietzsche

Socrates' enquiries usually lead to a damning conclusion: human beings are not as free or as knowledgeable as they tend to assume; their understanding of reality is not as solid or self-evident as it appears and the freedom it is thought to provide may well be illusory. This idea was first articulated in Plato's dialogues, but it soon acquired a wider historical dimension. Following Plato's lead, different authors and doctrines highlighted the inherent fragility of human cognition. Several explanations were offered as to how human ignorance limits human freedom and several methods were put forth to fix this problem.

This long and diverse history begins with Stoicism. For Chrysippus, Epictetus or Marcus Aurelius, freedom also requires a specific kind of education, leading to a complete redefinition of the aims and priorities that shape human life. By exposing and challenging the naïve objectivism of ordinary consciousness, the Stoics aim to become their own masters and to take their life in their own hands. Later, the idea that freedom requires a radical transformation of human cognition also played a central role in the evolution of Christian thought. But the deliverance promised by Christianity is not about independence or self-sufficiency. On the contrary, it can only be attained through the recognition of a higher form of authority, to which everything else is ultimately subordinated. Yet despite their lasting influence, Stoicism and Christianity are not the only heirs to Plato's founding idea. And though they will play an important role in the following analysis, its main focus will be placed elsewhere. I will pursue a wider historical thread and confront Plato's views with the work of two modern authors who have explicitly linked the notions of ignorance and truth to those of bondage and freedom—or slavery and mastery—and whose philosophical projects are also devoted to the double task of educating and liberating human consciousness. These authors are Georg Wilhelm Friedrich Hegel and Friedrich Nietzsche, and their ideas on the relationship between freedom and knowledge constitute, in my view, the most extreme historical developments of Plato's standpoint.

To justify this claim, let me start by outlining the main affinities shared by these three philosophical projects. For Hegel and Nietzsche, as well as for Plato, there is a fundamental disagreement between human beings' usual estimation of their cognitive powers and their actual value. On the one hand, most of us believe we have a relatively good grasp of reality. We are convinced we can discern what is at stake in most ordinary situations and make informed and efficacious decisions. On the other hand, upon closer inspection, most of our judgements are found to rest on arbitrary and unintelligible assumptions. Our usual understanding of reality is a matter of belief, rather than actual knowledge, and promises more than it can deliver.

For Plato, Hegel and Nietzsche, this problem is not merely theoretical. If we are indeed unable to vouch for the truthfulness of our judgements, our freedom may be seriously compromised. If our choices are based on unwarranted and unintelligible assumptions, their effects might prove unexpected, unfavourable or even harmful.

In these situations, although we may be free to choose what we please, we are unable to achieve what we want, or to discern what is in our best interest. Our choices are not empowering, but debilitating, and our freedom amounts to an involuntary form of imprisonment.

Finally, Plato, Hegel and Nietzsche all agree that this lack of freedom calls for a global reform of human cognition. For all three, it is up to philosophy to correct this problem, either by raising our usual standpoint to the standpoint of truth or by questioning the very standard of truth such a standpoint is supposed to meet. In either case, a philosophical reversal of human ignorance does not amount to a mere increase in the quantity of things we know, nor to the replacement of our usual beliefs with different, but equally uncertain or arbitrary beliefs. At stake in these three philosophical reforms is a *qualitative* transformation, aimed not at the content, but at the form of human cognition.

Despite these similarities, however, Hegel's and Nietzsche's accounts of the relationship between power, knowledge and freedom are much more than mere reiterations of Plato's. In the latter case, although Socrates insists on the need to break free from the chains of ignorance and opinion, it is at best unclear whether such release is actually possible. The dialogues are ambiguous with regard to their final purpose: all things considered, they can be read either as general descriptions of the imprisoned nature of human cognition or as actual attempts at liberation. In the first case, their task is largely negative. They represent the starting point of a long philosophical tradition, whose *critical* role is still shared, in its essence, by Kant's epistemology. In the second case, their task can be envisaged in a positive way, as an effort to move beyond the critical stage, towards an actual deliverance from ignorance.

In Hegel's and Nietzsche's case, there is no room for such ambiguity. Both authors are explicitly set on overcoming, albeit in very different ways, the stage of criticism, and both claim to have found an actual solution to Plato's challenge. In fact, the temerity of their projects is precisely what sets them apart from previous philosophical approaches. Their aim is not simply to expose the limitations of human cognition nor to detail the form and the scope of the imprisonment they produce. Rather, they take on the Promethean task of *setting humankind free*.

This introduction is not yet the place to consider how this transformation is to come about. Nonetheless, I will start by offering a brief overview of Plato's, Hegel's and Nietzsche's standpoints, so as to bring out their main similarities and differences.

1.3.1 *Plato's Philosophical Project*

Plato's characterization of human cognition is primarily based on the notion of εἶδος, or ἰδέα, usually translated as 'idea', 'form' or 'concept'. His first and most basic concern is to show that human beings' understanding of reality is entirely made up of εἶδη, or ἰδέαι. Ideas, forms or concepts are the basic currency of human

cognition, the shape in which everything whatsoever comes to be known, and their variety is virtually unlimited: they may stand for things, qualities of things, quantities, feelings, actions, logical functions, etc. This does not mean, however, that each cognitive insight corresponds to a different and unique idea. If that were the case, human experience would be purely nominalistic: every new insight would amount to a complete novelty and the world would be a permanent source of wonder. On the contrary, apart from maintaining that reason is entirely *eidetic*, Socrates points out that its structure is originally *synthetic*. Instead of ascribing a different εἶδος to each cognitive content, thoughts and judgements are grounded in a web of repeating and interchangeable εἶδη. As regards ‘all the ideas or forms, the same statement holds, that in itself each is one, but that by virtue of their communion with actions and bodies and with one another they present themselves everywhere, each as a multiplicity of aspects.’²⁶

But not all combinations of ideas are equally legitimate. And neither are they equally essential. Regarding the first of these observations, Plato compares the multiplicity of εἶδη to the letters of the alphabet: while some letters can be joined together to form words and sentences, others cannot; likewise, while some ideas can be combined in intelligible and coherent ways, others cannot.²⁷ Human cognition is bound to a specific grammar, without which reason and meaning are impossible. Secondly, regarding the essentiality of these combinations, Plato points out that the alphabet of human cognition is a *vertical* alphabet, divided into different hierarchical stages. It is structured as a vast conceptual tree, where each particular concept is both supported by the combination of lower-order ideas and subordinated to higher orders of meaning. To borrow an example from the *Republic* [596a], the terms ‘couch’ and ‘table’ refer to single, self-contained objects. Nonetheless, couches and tables are only recognized as such because they occupy a specific position within a wider conceptual hierarchy. Looking downwards, they are the genera or classes (γέννη) of which every particular couch or table partakes; looking upwards, they themselves partake of higher-order genera such as ‘object’ or ‘body’, and of universal genera such as ‘identity’ and ‘difference’, ‘rest’ and ‘movement’, ‘simultaneity’ and ‘duration’, etc. Naturally, this complex architectural structure is not usually apparent when one looks at a couch or lies on it. In everyday life, the notion of couch functions mostly as a ‘non-thematized assumption’, used as the basis for various kinds of judgements and actions.²⁸ The point, though, is that the intelligibility of this notion hangs on the simultaneous, albeit unconscious acknowledgement of the entire conceptual apparatus outlined above. If a specific couch were not regarded as an

²⁶ *Re* 476a5–7. The term εἶδη is usually employed by Plato in connection with ‘abstract’ notions such as unity, good, beauty, and the like. This does not mean, however, that it only applies to a specific part of human cognition. Plato favours these basic notions, which he arrives at by deconstructing more complex thoughts and ideas, because they are the cognitive bricks with which *everything whatsoever* is built. This issue is brought up in *Pa* 130c–e, where it is suggested that even the most ‘ridiculous’ or ‘worthless’ things, such as hair or mud, have their own εἶδη.

²⁷ See *So* 252e9ff. or *Phi* 17a8ff.

²⁸ Wieland, *Platon und die Formen des Wissens*, 149

individual instantiation of the form or genus ‘couch’—and hence as a solid object, with a determinate length, subsisting through time, etc.—it would not be recognized as it usually is, nor used to lie on.

Moreover, a third and essential aspect of Plato’s characterization relates to the practical nature of human cognition. As argued, for example, in the *Philebus* or the *Gorgias*, human beings’ usual account of reality is not exclusively, nor indeed primarily, motivated by theoretical concerns. In this respect, Plato anticipates a central element of the Stoic tradition: since men and women are never impartial or disinterested spectators of reality, knowledge is rarely, if ever, pursued *for its own sake*. Human beings wish to know things in order to achieve certain goals or avoid certain obstacles. Their understanding of reality is guided by a global diagnosis of what is *at stake* in their life—of who they are and who they want to become, of what is in their best interest and what is not, of what is worth pursuing and what can be neglected, etc. In Plato’s vocabulary, every human life is globally oriented towards what is perceived, at each given moment, as the *best* or the *greatest good* available (τὸ μέγιστον ἀγαθόν, or τὸ βέλτιστον).²⁹ And this fundamental drive extends to reason’s hierarchical system. Again, when sitting on a couch, one is usually not absorbed by such difficult questions as who one wants to become or what matters most in life. These interrogations are reserved to special, introspective moments, whose seriousness contrasts with the usual bustle of everyday life. Yet Plato argues that one’s usual conception of couches, tables and indeed anything else presupposes a specific, if inexplicit and provisional answer to these fundamental questions. Couches and tables are never perceived or defined on their own account, but as part of a wider existential narrative. Whereas for most people a couch is merely something to sit on, or a means of fulfilling other, more pressing tasks, for couch designers or sellers the εἶδος ‘couch’ is less generic. Since their life project is more directly tied to what a couch is or looks like, their interest in couches is more demanding. They are aware of certain characteristics of couches that most people tend to ignore, like their inner structure or their fabrication process.

These and other issues will be explored below. For now, let us summarize the main features of Plato’s portrayal of human cognition. We have seen, so far, that human thoughts and judgements are supported by a universal, stratified and self-referential web of ideas or concepts. Each mental representation, however simple, is already endowed with a specific alphabet, a specific morphology and a specific syntax. Each is already committed to a highly complex constellation of εἶδη, subordinated to a specific assessment of the structure and meaning of human life. But apart from unveiling this complex cognitive structure, Plato’s dialogues also aim to highlight how far it usually is from most people’s minds. Unlike Socrates, most human beings are usually blind to reason’s inner workings. Their grasp of the value and scope of their knowledge is very superficial and rarely called into question.

²⁹ See *Go* 468a–b, 499e8–9; *Sy* 205e7–206a1; *Me* 77b–78b; *Eu* 282a2; *Phi* 20d.

In everyday life, cognition as such is not really an issue. And this means, firstly, that most concepts are not experienced as concepts, but as the very things whose meaning they are supposed to convey. One's ordinary conception of couches and tables tends to coincide with couches and tables themselves and the very idea of cognitive access is thus disregarded. This also means, furthermore, that most concepts are automatically regarded as clear and coherent. One usually knows what is meant by the terms 'couch' or 'table' and has no difficulty employing them in different contexts.

These two assumptions are extensive not only to most human ideas, but also to all the other cognitive properties listed above. The fact that most concepts are able to partake in different conceptual arrangements is hardly construed as noteworthy or problematic. Likewise, the fact that most concepts can be decomposed into lower-order concepts, or joined together to form higher-order concepts, is also accepted as natural and self-evident. Ordinary consciousness breezes through these difficulties because it does not usually stop to consider them. Its relationship with concepts and conceptual arrangements is brief and inattentive, focused on practical rather than theoretical concerns. Socrates' enquiries, on the other hand, aim to highlight and reverse this natural tendency. They mean to show that the edifice of human reason, when examined more carefully, is anything but simple or self-evident.

Plato's usual method is to isolate a given concept or set of concepts and to challenge both their adequacy and their intelligibility. As regards the first of these operations, his strategy is to show that the meaning of any given concept is fully dependent on other concepts, whose definition is provided in turn by other concepts. The assumption that one's ordinary account of reality coincides with reality itself is therefore a fully *conceptual* assumption, backed by a constellation of concepts that are themselves supported by the same immediate assumption. Conceptual truth is usually a mere *presupposition* of truth, which is ultimately unable to vouch for its own validity. As regards the issue of intelligibility, Plato's argument is usually built along the same lines. Since the definition of a given concept is provided by a specific conceptual arrangement, couches and tables, for example, are defined by their participation in continuity, solidity, length, and so on. Yet these notions are also used to define many other objects and events. What makes them specific features of *this couch* is their participation in the concept 'couch'—but the latter is what it is only insofar as it refers back to continuity, solidity or length! This circularity is bound to arise because the very notions of participation, subordination and interchangeability, on which the entire alphabet of cognition is founded, are ultimately confusing and unintelligible.³⁰

By encouraging consciousness to turn inwards and confront its own ignorance, Plato plants the seed of a terrible suspicion. If human cognition is based on a vast web of assumptions, whose truth-value rests on further, equally unverified

³⁰See *So* 253b8ff.: What human beings usually lack is 'a clear perception of one form or idea extending entirely through many individuals each of which lies apart, and of many forms differing from one another but included in one greater form.'

assumptions, reason as we know it may well amount to a gigantic misunderstanding. Indeed, the scope of Plato's critique should not be minimized: it refers not to a specific domain of reason, nor to a specific brand of ideas or concepts; his denouncement is primarily aimed at the very form of human reason, and affects therefore its entire cognitive repertoire. The comparison between the human race and a community of prisoners, between human life and an involuntary dream,³¹ between human consciousness and a diseased or deranged organ³²—all of these images refer to an all-pervasive problem. They suggest that the entire edifice of human ideas is nothing more than a highly complex patchwork of beliefs, with no clear foundation. Human cognition, in its habitual form, may simply reflect 'the aberration of a soul that aims at truth, when the understanding passes beside the mark.'³³

If Plato's worries are justified, this problem is liable to have serious practical consequences. If the concepts underlying our daily judgements are indeed unfounded and unintelligible, the vital narrative that guides our thoughts and actions is also open to attack. Apart from couches and tables, or trees and cars, Plato's denouncement also extends to our usual idea of what is good or bad, pleasurable or painful, just or unjust, etc. What is ultimately being challenged, therefore, is the overall cogency of our usual diagnosis of who we are and where we stand, of what is in our best interest, of what is worth pursuing, etc. In short, what is at stake is the actual freedom to steer the course of our lives—the sovereignty we take for granted, but are unable to guarantee.

For Plato, the only way to deal with this problem is to do as Socrates does, namely to bring reason's alphabet into focus and determine which ideas are in play in one's usual standpoint, how they relate to one another and how reliable they really are. This ambitious task requires a specific science, and indeed 'the greatest of sciences',³⁴ namely dialectics (ἡ ἐπιστήμη διαλεκτική). In contrast to the captivity of ordinary consciousness, dialectics is 'the science that belongs to free men'³⁵ and the key to a philosophical reform of human cognition. But since most human beliefs are usually assumed to be true, this reform is at first merely negative. The philosopher's first task is to free men and women from their 'conceit of knowledge', to turn their self-confidence into perplexity and to clear the way for a new form of cognition.

³¹ *Re* 476c–d, 520c–d, 533b–c, 534c–d; or *Th* 158b–d. See Sect. 11.1.

³² *So* 228c10–d2

³³ *So* 228d1–2. Plato describes this peculiar kind of delusion as a form of παραφροσύνη, i.e. madness. The prefix παρα- is used in this context to convey both the idea of physical deflection or deviation—the soul passes *beside* truth just as an arrow passes *beside* the mark—and the idea of mental or spiritual deviation — the idea of madness or derangement. The term παραφροσύνη is often used in the *corpus hippocraticum*, where it is usually contrasted with σωφροσύνη, that is, with the health or sanity required to see things *as they really are*. In the *Sophist*, as in many other Platonic texts, human reason's alleged health is exposed as a form of disease in disguise. Men and women do not know what they think they know, nor do they understand what they take to be familiar and self-evident.

³⁴ *So* 253c4–5

³⁵ *So* 253c7–8

Those who are willing to undergo this transformation ‘will have to change the opinions which they had at first accepted, so that what was great will appear small and what was easy, difficult, and all the apparent truths in arguments will be turned topsy-turvy.’³⁶

1.3.2 *Hegel’s Phenomenological Project*

The idea that human life, as regards one’s usual powers of cognition, is not a free life has wielded great influence throughout the history of Western thought. Yet few authors have taken Plato’s insistence on the need for a philosophical liberation of humankind as seriously and as far as Hegel. His critique of human cognition represents in many ways an amplified and systematic reworking of Plato’s critique. As one commentator observed, Hegel’s ‘whole philosophical system, in all its incredible breadth and detail, can be understood as a single extended demonstration of the importance and meaning of freedom. Moreover, Hegel’s philosophy is not only *about* freedom, but also claims to be productive *of* it: in the course of his philosophical investigation of what it means to be free, Hegel arrives at the view that freedom depends upon the practice of philosophy.’³⁷

Like Plato’s views on freedom and philosophy, scattered throughout different dialogues, Hegel’s treatment of this issue is not confined to a single text. The concept of freedom is taken up at different junctures of his philosophical system, corresponding to different degrees of cognitive insight, or different ‘shapes of Spirit’. Furthermore, like Plato, Hegel views philosophy as an elevation to the standpoint of science (*Wissenschaft*, ἐπιστήμη), i.e. to a true and clear understanding of reality.³⁸ Therefore, depending on the distance separating one’s cognitive standpoint from a truly epistemic or scientific standpoint, freedom can assume different forms and modes of expression.

Hegel’s *Outlines of the Philosophy of Right*, devoted to the ‘development of the idea of the absolutely free will’,³⁹ is commonly regarded as the most complete and direct exposition of his views on freedom⁴⁰—from the more immediate notions of free will and free choice to the social and political institutions required for the actualization of freedom. However, as Hegel points out in the preface, the dialectical

³⁶ So 234d6–e1

³⁷ Dudley, *Hegel, Nietzsche and Philosophy*, 15

³⁸ In Hegel’s writings, the term ‘science’ (*Wissenschaft*), like the Greek term ἐπιστήμη, is usually a synonym for actual knowledge, as opposed to mere beliefs or opinions. It must not be confounded, therefore, with the modern word ‘science’, which refers to an organized body of knowledge on a specific subject.

³⁹ PR, 50 / HW 7, 87

⁴⁰ Along with the *Encyclopaedia*’s section on Objective Spirit, in many ways a preliminary version of the *Philosophy of Right*.

progression developed throughout this work presupposes an acquaintance with the ‘philosophical’ way of knowing already set forth in his logical writings, namely in the *Science of Logic*.⁴¹ There, in turn, the reader is warned in the introduction that the text’s starting point does not amount to an absolute beginning, as it were, since the ensuing developments already take place at the level of scientific thought—that is, at a level which is essentially different from one’s usual, ‘pre-philosophical’ or ‘pre-scientific’ standpoint.⁴²

Accordingly, the first major step towards the philosophical release of ordinary consciousness must be looked for not in the higher stages of Hegel’s system, but in an earlier propaedeutic work entitled *Phenomenology of Spirit*. Its overall aim is summed up in the ad published by Hegel in 1807 in Jena’s *Allgemeine Literatur-Zeitung*: ‘The *Phenomenology of Spirit* aims to replace the psychological explanations and abstract accounts of the foundation of knowledge. It considers the *preparation* to science from a point of view that makes it a new, interesting, and indeed the first science of philosophy.’⁴³

The following analysis will focus mainly on Hegel’s *Phenomenology*. The debate concerning his later assessment of this work, as well as its articulation with his mature philosophical output, will be bracketed off in favour of a more direct approach to his phenomenological method. In keeping with my focus on the notion of paideia and its various historical developments, Hegel’s views will come up in connection with his unique understanding of the relationship between freedom, knowledge and education—a relationship that is explicitly defined and developed throughout the *Phenomenology*. Like Plato’s educational project, and unlike the Hegelian system’s subsequent stages, the *Phenomenology* does not require a previous philosophical education. Its starting point is human consciousness in its native or ordinary form, with all the assumptions that shape its usual outlook on reality.

Hegel’s phenomenological progression is set in motion by the realization that the knowledge held by ordinary consciousness is not actual knowledge, but only the assumption or the *appearance* of knowledge.⁴⁴ Like Plato, he aims to expose and overcome the spuriousness of ordinary concepts and ideas through an immanent critique of human cognition, designed to ‘[lead] the individual from his uneducated standpoint to [the standpoint of] knowledge.’⁴⁵ Moreover, Hegel also believes that a higher cognitive standpoint entails a higher degree of freedom. The phenomenological consciousness echoes the imprisoned consciousness of the *Republic*’s cave

⁴¹ See PR, 4 / HW 7, 12

⁴² See SL, 28f. / HW 5, 17f.

⁴³ HW 3, 593

⁴⁴ Hegel’s emphasis on the ‘appearance’ of knowledge is deliberately ambiguous: the phenomenal knowledge that is taken up and criticized in the *Phenomenology* is both the *knowledge of appearances* (viz. of that which immediately appears or comes into view: *Erscheinung*, φαινόμενον) and the *appearance of knowledge* (a mere semblance or pretence: *Schein*). As we shall see, the transition from the first to the second of these meanings is the *principium motionis* of the entire phenomenological progression.

⁴⁵ PS, 16 / HW 3, 31

dwellers, and its journey mirrors that of Plato's allegory, where ignorance and darkness are eventually 'cut short' 'by a sunburst which, in one flash, illuminates the features of the new world.'⁴⁶

In particular, Hegel explores an aspect of Plato's narrative whose importance is often disregarded. As we will see below, the allegory of the cave is not about a simple contrast between ignorance and knowledge, or captivity and freedom, but about an entire cognitive scale—a *scale of freedom*, as it were, with several intermediate stages. Initially, as one of the prisoners is released from his chains, his ignorance gives way to a new and more sophisticated standpoint.⁴⁷ His idea of reality is no longer reduced to the shadows on the wall and his life-long captivity is revealed as such. But this discovery is only the first in a long series of transformations.⁴⁸ Once the prisoner overcomes his *physical* bondage, he is led into a new form of imprisonment, namely an *optical* imprisonment, imposed by the fire's blinding glare. Afterwards, as he regains his eyesight, makes out the cave's ascending path and reaches the cave's entrance, the sun's brightness prevents him once again from seeing ahead. To avoid this new obstacle, he turns his eyes to the shadows cast by objects, then to the reflections of objects in water, and only then—eventually—to real objects.

This long narrative brings out just how difficult it is to shed one's usual cognitive skin and embrace a fully epistemic standpoint. And yet the progression towards freedom and truth is not experienced as a progression, but as a succession of unique and global transformations, whose final outcome is systematically postponed. The initial acknowledgement that the world is irreducible to the shadows on the wall leads to the discovery of a new world, hastily recognized as the real world. But just as the shadows turned out to be mere shadows, this new world is eventually revealed as a mere cave. What seemed final is once again provisional and the progression's ending is shown to be a new beginning. The same illusion is then transferred to the shadows outside, and again to the reflections on water. Until the very end, the prisoner is led from one world to the next, subjected to increasingly sophisticated forms of deception.

A similar mechanism is deployed throughout the *Phenomenology*, where the abyss separating ignorance from knowledge is also filled by a succession of intermediate stages. Starting from consciousness' most immediate cognitive outlook, Hegel draws up a catalogue of the different kinds of deception encountered along the 'pathway to Science'. In each stage, consciousness is led to confront its own ideas and assumptions and to test their solidity. This exercise leads to the discovery that the knowledge it takes for granted is inherently contradictory. As in Plato's account,

⁴⁶PS, 7 / HW 3, 18 f. See Philonenko, *Leçons Platoniciennes*, 331, and *Lecture de la 'Phénoménologie'*, 24f.; Stekeler-Weithofer, *Hegels Phänomenologie des Geistes*, 539.

⁴⁷Here and in future references to the allegory of the cave, I follow Plato's use of masculine pronouns to refer to the prisoners. This does not mean, of course, that the allegory applies only to men. It is meant to illustrate the cognitive situation of each and every human being.

⁴⁸See *Re* 515ff.

this cognitive volte-face entails the collapse of an entire worldview, but also the revelation of a more sophisticated one. Having revised its cognitive outlook, consciousness puts forward a new version of reality, based on a new set of principles. What is more, it embraces a new existential narrative, shaped by new aims and priorities. But this change is still only provisional: once these new ideas are subjected to a closer examination, they turn out to be mere beliefs; and once the latter are taken up and examined, they are also found lacking. This procedure is repeated over and over again, at increasingly higher cognitive stages, until no assumption or illusion is left to uncover.

The progression's final stage is no longer a mere version of reality, but 'the actual cognition of what truly is.'⁴⁹ The Platonic and Hegelian projects are both aimed at this final release, but their strategies are very different. Plato's approach is diverse and episodic. It combines a series of programmatic indications about the kind of reform required for the attainment of truth and a series of practical examples of how they should be applied. Overall, however, the dialogues amount to the pieces of an unsolved puzzle, which is up to the readers to interpret and complete. Hegel's approach, on the other hand, is global and systematic. His work is not merely an introduction to philosophy or dialectics, but a complete and detailed account of how to go about transforming one's usual standpoint into an educated standpoint. While Plato points the way to freedom, the *Phenomenology* is itself a pathway of liberation, leading from the immediacy of 'sense-certainty' all the way to the freedom of 'absolute knowing'.⁵⁰

The entire *Phenomenology* is built on the conviction that the realm of illusion inhabited by ordinary consciousness is not endless or chaotic. Hegel's phenomenological catalogue is not a random catalogue, but a *systematic* one, with each stage proceeding from the previous one and necessarily bound to the next. Furthermore, since this catalogue can only be brought to an end when every illusion has been exposed and refuted, it is also a *global* catalogue, including every major mode of deception discoverable throughout human life. Hegel undertakes nothing less than the actual inventory of all the main worldviews available to human consciousness, as well as all of their combinations. His aim is to secure the attainment of truth by exhausting every form of untruth. And his progression is therefore an encyclopaedic progression, extending to every domain of human existence: from logic to language, from the natural to the social sciences, from ethics to aesthetics.

⁴⁹PS, 46 / HW 3, 68

⁵⁰As noted earlier, Hegel's engagement with freedom is also a central concern in his later, systematic works, which raises the issue of whether the freedom attained at the end of the *Phenomenology* is already a complete or absolute freedom—in which case Hegel's system is already directed at a free, educated consciousness—or whether it is still a new and more sophisticated form of imprisonment—to be overcome at a later stage. This issue leads to the wider controversy regarding the role assigned by Hegel to the *Phenomenology*, which cannot be dealt with here. My present purpose is to contrast Plato's, Hegel's and Nietzsche's methodological approaches to the relationship between freedom and education. In Hegel's case, this theme is best illustrated in the *Phenomenology*, explicitly defined as the 'detailed history of the education of consciousness itself to the standpoint of Science.' (PS, 50 / HW 3, 73)

1.3.3 Nietzsche's Metacognitive Project

Like Plato and Hegel, Nietzsche advocates a profound reform of human cognition. In line with the imagery employed by his predecessors, he compares human beings to 'involuntary slaves',⁵¹ or 'fettered spirits', 'chained . . . to [their] pillar and corner'.⁵² Descriptions of this kind abound in his writings, as do prophecies of a 'great liberation', or appeals to the 'dangerous curiosity for an undiscovered world', ruled by 'the master's privilege of the free spirit'.⁵³ But although this kind of language is indeed evocative of Plato and Hegel, Nietzsche opposes the Platonic and Hegelian solutions to the problem of freedom and calls for an altogether different kind of release.

The key to understanding this difference lies once more in Plato's writings, and specifically in his insistence on the fact that human beings are never indifferent to the truth-value of their thoughts and judgements. In the course of our lives, we usually assume our most basic beliefs to be true and tend to resist any suggestion to the contrary. Our ideas about ourselves and the surrounding world are not conjectural, but assertive, and indeed dogmatic. We take things to be as we believe them to be and not otherwise, and it is based on this certainty that we interpret what goes on around us, make specific choices and act in specific ways. Moreover, even when we do think of something in conjectural terms, we think of it nonetheless as being in a certain way rather than another. Our understanding of it presupposes a specific version of things, even if only momentarily maintained.

In short, our usual representation of reality entails the simultaneous representation of its truthfulness—which requires, in turn, the simultaneous representation of the notion of truth. Hegel, in the introduction to the *Phenomenology of Spirit*, makes this point in the following terms: 'Consciousness simultaneously *distinguishes* itself from something and at the same time *relates* itself to it'.⁵⁴ More exactly, it distinguishes a *being-for-another*, i.e. its knowledge of reality, from a *being-in-itself*, i.e. reality as such. Hence, 'whatever is related to knowledge or knowing is also distinguished from it, and posited as existing outside of this relationship; this *being-in-itself* is called *truth*'.⁵⁵

The notion of truth is the silent backdrop against which all human thoughts are measured. And this is why, for Plato, humankind as a whole, 'regarding both education and the lack of it', is trapped inside a cave. Indeed, it could just as well be the case that the truthfulness of our representation of reality were a matter of indifference to us and that our lives were spent in a sort of logical relativism, which would render every representation as true as the next one, or all of them true in their own way. In that case, freedom and knowledge would not be as closely connected

⁵¹ KSA 11, 75

⁵² HAH, 6 / KSA 2, 16

⁵³ HAH, 7 f. / KSA 2, 16ff.

⁵⁴ PS, 52 / HW 3, 76

⁵⁵ PS, 52f. / HW 3, 76

and it would make no sense to speak of a cave, or shackles, or prisoners in the dramatic way Socrates does. But Plato's allegory aims to show that this is never the case: human beings are imprisoned because their representation of reality falls short of reality itself. There is an original and potentially insurmountable opposition between the cave's darkness and the sunlit world supposed to exist outside.

This fundamental tension must not be reduced to a simple antithesis, in which truth as such is placed on one side, as an independent or preformed ideal, and ignorance on the other. Plato's point is more subtle and much more serious: he does not argue simply that truth exists in its own right, in some celestial kingdom beyond human life; rather, he strives to show that all human beings, insofar as they think anything whatsoever, are bound to posit the notion of truth ahead of themselves, as it were, and to conceive their understanding of reality as either truthful or falling outside of truth. In this broader sense, truth is just as much a predicate of human thoughts and judgements as it is their underlying condition.

Now the main bone of contention between Nietzsche and the Platonic tradition is precisely the notion of truth, and particularly the referential status it usually enjoys. For Plato, Hegel and a vast lineage of thinkers, philosophy's starting point is the disagreement between human consciousness' usual account of reality and a true account of reality. Our lack of freedom, our lack of power, our inability to attain self-fulfilment are all secondary symptoms of a more basic cognitive predicament. Human life, down to its most essential aspects, is dependent on a global expectation of truthfulness, but the latter is continually thwarted by reason's inherent contradictions.

Plato and Hegel agree on this general diagnosis but differ in their methodological approach. With Nietzsche, on the contrary, this common view no longer applies. To be sure, he also believes humankind to be enslaved and in need of liberation; moreover, he also takes this lack of freedom to be primarily caused by consciousness' incapacity to meet a predetermined standard of truth. But whereas Plato and Hegel take truth as such for granted, and focus on how to attain it, Nietzsche takes a step back and questions the very need to think and choose in terms of truth and falsehood. While Plato describes human life as being 'aimed at truth' but passing 'beside the mark', Nietzsche challenges the very existence of such a mark, as well the obligation to aim at it. In his eyes, the concern with truth is itself one of the dogmatic presuppositions consciousness hangs on to in everyday life. A serious philosophical critique, he argues, must extend to the very criterion in which every critique is usually grounded, namely the notion of truth, and expose its uncritical nature.

For Nietzsche, life is not primarily about truth, but about power and vitality. Men and women are indeed concerned with satisfaction and self-fulfilment, but these imperatives are not intrinsically cognitive, let alone rational. Cognition itself is a symptom of an inner volitional core whose immediacy predates the need to distinguish between true and false ideas, accurate and inaccurate perceptions, rational and irrational judgements, etc. Human will is originally free and unhindered, open to an endless repertoire of possibilities and modes of expression. In everyday life, however, it is reduced to a dogmatic concern with truth—a 'will to truth' that outweighs

all other forms of valuation. This artificial exclusivity represents, for Nietzsche, a form of imprisonment. By sacrificing the richness of its resources to a single cognitive standard, the will is reduced to an anaemic or slavish version of itself. Accordingly, only a standpoint released from its cognitive aspirations—a *metacognitive* or *metagnoseological* standpoint, which has traded its usual concern with truth with a more primordial concern with life and vitality—will enable a free or independent existence.

Nietzsche's exploration of this theme is not uniform. His critique of truth and rationality is grounded in diverse and sometimes conflicting arguments, borrowed from the domains of epistemology, biology, history and aesthetics. Overall, however, his views on the relationship between truth and freedom fall into two main lines of enquiry. At times, he seems to regard the 'will to truth' as a natural cognitive tendency, inseparable from the natural use of reason. If all human concepts are indeed supported by the notion of truth, human cognition, in its usual form, is globally imprisoned. In light of this acknowledgement, Nietzsche insists on the wideness of the gap separating natural consciousness from a truly free consciousness and highlights the depth of the transformation required to bridge this gap. The deliverance he advocates entails a full ontological revolution: the crossing of an abyss from which 'one returns *newborn*, having shed one's skin.'⁵⁶

On other occasions, Nietzsche puts the emphasis on the historical nature of this transformation. He suggests that the 'will to truth' is not an essential feature of human consciousness, but the result of a long cultural process, leading to the universal dissemination of an artificial cognitive attitude—the 'philosophical', or 'metaphysical', or 'theological' tendency to reduce reality to the dialectical opposition between truth and untruth, right and wrong, good and evil. Revisiting the history of human reason, Nietzsche contrasts the primitive wisdom of the ancient world (σοφία) with the scientific attitude of the classical and modern worlds (ἐπιστήμη).⁵⁷ Primitive human beings, he argues, favoured life over cognition; they were free, like children or artists, to create their own values. Modern human beings, on the contrary, take truth and reason too seriously; forgetting themselves and their creative powers, they place truth at the centre of their lives.

The first great step was taken by Socrates, the promoter of 'a new, unheard-of esteem for knowledge and insight.'⁵⁸ He is the founder of a new cultural type, the 'theoretical man', to whom freedom and happiness are essentially a matter of knowledge. Unlike others before him, Socrates aims to explain away life's contradictions and to render the world a thoroughly rational place. He equates the True, the Good and the Beautiful and turns truth simultaneously into a scientific, moral and aesthetic imperative. This attitude paved the way for a new and even greater form of imprisonment, brought about by Judaism and Christianity. The Jews and the Christians inherited Socrates' and Plato's fixation with the divine value of truth and

⁵⁶GS, 7 / KSA 3, 351

⁵⁷See KSA 7, 448.

⁵⁸BT, 65 / KSA 1, 89

developed a slavish mistrust of strength, power and the concrete pleasures of life. This ‘denial of life’, coupled with the epistemological and political misconceptions of modern European philosophy, led to the nihilistic ‘disease’ which has ‘spread unevenly across Europe.’⁵⁹

Nietzsche’s point is not only that these transformations have shaped world history, but also that they have reshaped human consciousness, conditioning the way most men and women view themselves, each other and the surrounding world. Strongly opposed to this legacy, Nietzsche advocates a major cognitive change—a ‘transvaluation’ of modern values, whose effects are personified by a series of mythological anti-heroes: Dionysus, the Greek god of music and drunkenness; Zarathustra, the rebellious hermit; and the Antichrist. All of these figures embody, in different ways, the recuperation of a primitive, pre-modern ideal, averse to the laws of cognition and rationality. But since this primordial worldview has grown distant and unintelligible, these figures also personify an entirely new standpoint, free from the dialectical opposition between truth and untruth.

Although Nietzsche’s project raises a number of difficult questions, with regard to its scope, meaning and viability, its contrast with the two previous projects is now more apparent. Following Plato and Hegel, he views the problem of freedom as ‘the problem of science itself.’⁶⁰ Nonetheless, his definition of freedom is incompatible with the scientific ideal espoused by Plato and Hegel. For Nietzsche, what keeps humanity in chains is not the lack of a scientific mode of cognition, but the yearning for it. Freedom is not achieved by overcoming ignorance, but by overcoming the very need to do so.

1.4 Historical Affinities

Hegel and Nietzsche disagree on the merits of Plato’s philosophical project, but both acknowledge its lasting historical influence. Furthermore, both emphasize the transitional nature of their own historical age and the importance of the role played by their own philosophical standpoint in the global narrative initiated by Plato. Indeed, both Hegel and Nietzsche view their solution to the ‘problem of science’ as the necessary culmination of a long historical progression, whose different stages are regarded as either positive or negative achievements depending on whether they lead human consciousness towards or away from the kind of deliverance they prescribe.

In Hegel’s account, freedom is dependent on a cognitive progression, leading from an immediate to an educated outlook on reality. This progression is not just logical or epistemological: it is meant to capture the extraordinary complexity of human life, in all its dimensions. At stake in each stage of development is a global worldview, with logical and epistemological, but also psychological, moral and

⁵⁹BGE, 101 / KSA 5, 139

⁶⁰BT, 4 / KSA 1, 13

political implications. Moreover, these different scenarios are not purely conceptual or imaginary. They also reflect, albeit in an indirect or idealized way, consciousness' concrete historical evolution. The whole *Phenomenology* is built on the idea, central to Hegel's entire philosophy, that there is a necessary link between reason's dialectical development and its actual historical development—in other words, that the history of human thought, despite its share of contingency and randomness, amounts to a global movement away from ignorance and towards more sophisticated forms of rationality. Hegel thus regards modern reason as the necessary, albeit indirect outcome of Greek philosophy; and Plato's philosophical project as the starting point of a long and complex philosophical saga, leading all the way to German Idealism.

For Hegel, Kant's critical project amounts in many ways to a systematic reform of Plato's critique of cognition. In his words, Kant 'turns back to the standpoint of Socrates', but his philosophy is 'placed on a higher level.'⁶¹ Nevertheless, although his standpoint is closer to the epistemic standpoint sought by Plato and his successors, his conclusions are far from reassuring. Instead of solving the 'problem of science', Kant's criticism seems to aggravate it by revealing just how problematic it really is. His insistence on the fundamental discrepancy between human cognition and the actual things it is meant to cognize—between the appearance of reality and reality in itself⁶²—is not yet liberating, in the Platonic or Hegelian sense of the word. Rather, it offers a new and even more damning image of human beings' natural servitude.

The issues of whether and how Kant's cognitive dualism can be overcome are a constant theme in Hegel's early writings and constitute the immediate philosophical ground on which the *Phenomenology of Spirit* is erected. While Kant does highlight the reciprocal implication of reason and reality, his philosophy remains trapped within this antithesis: 'it makes the identity of the opposites into the absolute terminus of philosophy, the pure boundary which is nothing but the negation of philosophy.'⁶³ Yet 'the sole Idea that has reality and true objectivity for philosophy is the absolute suspendedness of the antithesis. This absolute identity is not a universal subjective postulate never to be realized. It is the only authentic reality.'⁶⁴ Therefore, if philosophy is to become more than a mere critique of human cognition, it must be able to unite both sides of Kant's equation and provide consciousness with a standpoint that is both subjectively valid and objectively true.

Fichte's and Schelling's efforts to bring about such unification amount to new and more sophisticated stages in the path towards science. Yet both of their solutions replicate Kant's dualism by replacing his 'universal subjective postulate' with their own subjective postulates. By professing the unity of thought and reality and attempting thereby to seal the breach between self-consciousness and objectivity,

⁶¹ LHP, vol. 3, 424 / HW 20, 329f.

⁶² See, for ex., *Critique of Pure Reason* A114, B311f.

⁶³ FK, 67 / HW 2, 302

⁶⁴ FK, 68 / HW 2, 302

Fichte's 'I' and Schelling's 'intellectual intuition' are ultimately just as undialectical as Kant's critical approach. Whereas the former postulates the formal antithesis between the self and its object, the latter postulates the formal unity of these two instances. In so doing, Hegel argues, both solutions take as starting point that which can only be known as a *result*, and maintain the presuppositive element that a true philosophical education must be able to overcome.⁶⁵

As Hegel writes, truth is not 'an *original* or *immediate* unity as such', but a 'self-moving sameness', or a 'reflection within itself'.⁶⁶ If philosophy is to result in actual knowledge, it must cease to envisage the opposition between truth and appearance as either inclusive or exclusive. It must move beyond the immediate contrast between self-sameness and difference and acknowledge contradiction itself as the driving force behind reason and reality. In this new conceptual framework, subjectivity and objectivity no longer stand for self-identical or self-excluding instances. They are defined instead in dynamic terms, as the movement through which each gives way to the other, thereby becoming what it is.

This dynamic mode of identity—'the identity of identity and non-identity'⁶⁷—is the founding principle of Hegel's logic and the cornerstone of his entire philosophical system. The revolution it sets in motion represents the culmination of a long conceptual journey and the beginning of a new philosophical era, dramatically announced in the *Phenomenology*:

Ours is a birth-time and a period of transition to a new era. Spirit has broken with the world it has hitherto inhabited and imagined, and is of a mind to submerge it in the past, and in the labour of its own transformation.⁶⁸

Interestingly, Nietzsche also insists on a major historical change. Half a century later, he also exalts his own age as a time of transformation; and like Hegel, he also points out the conceptual continuity between Plato, Kant and post-Kantianism. Yet he does not regard his own standpoint as the final stage in the historical and philosophical ascent towards science. On the contrary, insofar as the Western philosophical canon reflects Platonism's historical evolution, Nietzsche's intended liberation is also a break with history itself. He aims to recuperate a pre-historical and pre-scientific mode of cognition in which the Socratic dream of absolute knowledge plays as little part as possible.

To situate this transformation, it is useful to compare the historical roles assigned by Hegel and Nietzsche to Kant's *Critique of Pure Reason*. As Daniel Breazeale points out, 'both thinkers begin by acknowledging and affirming the crisis in

⁶⁵ Hegel's engagement with Kant and post-Kantian idealism is of course much more nuanced. My present aim is merely to highlight the structural affinities between Hegel's and Nietzsche's accounts of the historical evolution of Plato's philosophical project—both of which assign a central role to Kant's critical philosophy.

⁶⁶PS, 10 / HW 3, 23

⁶⁷DFS, 156 / HW 2, 96

⁶⁸PS, 6 / HW 3, 18

thinking brought about by the “all destroying” Kant.⁶⁹ In the first case, though Hegel acknowledges the specificity of the impasse revealed by Kant’s critical diagnosis, he regards it as the logical outcome of the ‘Socratic standpoint’, and hence as a necessary stage in the global progression towards science. For Nietzsche, on the other hand, Kant’s critique is not a prelude or a preparation to science, but the first major evidence of the failure of philosophy’s scientific aspirations. It is a premonition of the approaching break with the ‘old order of things’ and the restoration of an altogether different historical-philosophical paradigm, announced in a no less dramatic manner:

The catastrophe slumbering in the womb of theoretical culture is gradually beginning to frighten modern man; in other words, he is beginning to suspect the consequences of his own existence; . . . Meanwhile great natures with a bent for general problems have applied the tools of science itself, with incredible deliberation, to prove that all understanding, by its very nature, is limited and conditional, thereby rejecting decisively the claim of science to universal validity and universal goals.⁷⁰

The ‘great natures’ Nietzsche is referring to are Kant and Schopenhauer, whose scientific criticism of science prompted the crisis of modern reason and the need for a radical change of perspective. Kant’s critical philosophy, in particular, dealt a serious blow to the belief in the possibility of an actual communion between the subjective and objective realms—a belief underlying both the *natural* conviction that one’s thoughts provide an adequate perception of reality and the *philosophical* conviction that one’s usual standpoint, albeit lacking such perception, may come to acquire it by means of education. It must be noted, however, that Nietzsche’s enthusiasm for this kind of criticism was short-lived. As with many other of his philosophical infatuations, he soon came to regard Kant’s and Schopenhauer’s scepticism—as indeed the whole tradition of modern German philosophy—as a continuation rather than an actual break with the scientific tradition. By exposing the intrinsic limitations of human knowledge, Kant and his successors did not abandon the Platonic canon after all. They merely uncovered ‘a hidden path to the old ideal’.⁷¹ Once ‘reality was made into mere appearance’,⁷² a ‘real world’ was placed beyond it and accorded different names: a ‘Thing in itself’ (Kant), ‘the Absolute’ (Fichte and Schelling), ‘Spirit’ (Hegel), or ‘Will’ (Schopenhauer). But the operating principle in all of these concepts is still the old standard of truth.

In Nietzsche’s later writings, these different perspectives are increasingly treated as symptoms of the same fundamental problem, which he refers to, almost interchangeably, as the ‘idealist’, ‘metaphysical’, ‘moral’, ‘religious’ or ‘theological’ fallacy of modern reason. The idea of an immanent (and imminent) collapse of scientific reason, vaticinated in the *Birth of Tragedy*, is replaced by an increasingly

⁶⁹Breazeale, ‘The Hegel-Nietzsche Problem’, 163. The term ‘all destroying Kant’ refers to Moses Mendelssohn’s well-known comment about the *Critique of Pure Reason*.

⁷⁰BT, 87 / KSA 1, 117f.

⁷¹A, 9 / KSA 6, 176

⁷²A, 9 / KSA 6, 176f.

severe condemnation of philosophy in practically all of its forms. Nietzsche's efforts are devoted to unveiling a new, post-scientific and post-moral world, which remains for the most part unexplored. And although he does recognize, in past historical ages, a few promising premonitions of his own standpoint, he regards the latter as the first real stage in the path towards freedom.

1.5 Research Background

This book is divided into three main parts, centred on Plato's, Hegel's and Nietzsche's views on the philosophical release of human consciousness. As already noted, this issue is not confined to the work of these three authors. A detailed genealogy of the relationship between freedom, power and knowledge would have to include other authors and philosophical traditions. In what follows, however, I argue that these three philosophical projects represent three main landmarks in the debate we are considering, and the conceptual boundaries within which other theories and approaches operate.

Plato raised the issue. His dialogues highlight human reason's native imprisonment and the need for a full cognitive release. This idea gave rise to a long critical lineage, whose different members reiterate and amplify reason's intrinsic limitations. But if Plato was the starting point, Hegel and Nietzsche are, in a way, the ending—or rather, two alternative endings, pointing in two opposite directions. Hegel's project builds on Plato's standpoint and proposes to fulfil its main challenge. Nietzsche rejects Plato's standpoint and sets out to reverse its legacy.

This illustrious triangle encloses an entire range of possibilities, concerning one of philosophy's most important debates. Indeed, these three projects may be regarded as conceptual points of referral not only for the classical, medieval and modern understanding of freedom and truth, but also, to some extent, for post-modern and contemporary approaches to this relationship. Insofar as the works of Husserl, Heidegger or Wittgenstein, for example, highlight the cognitive dimension of freedom and power, they can also be located within this triangle and interpreted in light of its different alternatives. This does not mean, however, that this issue is exhausted by this kind of approach. There are, of course, other definitions of knowledge and other ways of thinking about freedom. It can be argued, in fact, that freedom is not as closely linked to cognition and truth as Plato and his philosophical heirs would have us believe. One of the central traits of contemporary philosophy is precisely its general mistrust of the classical notion of truth and the central role ascribed to reason and rationality in the Western philosophical canon. Many contemporary thinkers, in search of new ways of grounding freedom, happiness and self-fulfilment, sought to shift the emphasis placed on the contrast between truth and error to other forms of valuation, linked to the affective, psychic, cultural, social or even sexual dimensions of human life.

It may also be argued that the Platonic tradition, with its fixation on epistemic truth, failed to appreciate the different modes of truthfulness and struggle against deceit that finite human beings are capable of to various degrees. Indeed, the idea of a global release of human consciousness does not seem to take into account, much less value, the different ways in which local forms of knowledge and insight can meet the standards relevant to them. Instead of one truth, leading to one kind of freedom, there may well be several kinds of truth, each of them valid and liberating in its own way.

This kind of criticism is not without merit. However, in defence of the Platonic approach, a brief clarification is in order. Plato's main goal is indeed to raise the finite human standpoint to the standpoint of truth, but this transformation is not merely quantitative: that is to say, his project is not simply about the transition from a consciousness that knows too little to one that knows a lot—or to one that knows everything. Were that the case, freedom would indeed be the prerogative of an omniscient, god-like standpoint, infinitely distant from ours. In my view, however, Plato's project should be understood in a chiefly qualitative way: if his suspicions are right, the problem with human beings is not so much that we do not know enough, but that we ignore what we claim to know—and even what we claim to know best. In other words, the trouble is not that we do not think correctly, as it were, but that we *have yet to start thinking*. Plato's indictment of ordinary consciousness is thus much more radical than might be supposed: he suggests that *all* forms of insight, whether regional or universal, theoretical or practical, may turn out to be the contrary of what they seem, namely forms of ignorance in disguise.

Of course we can oppose Plato's suggestions and maintain that human beings are not as ignorant as he claims, or that human reason is not as open to delusion as he seems to believe. But if we take his criticism seriously, as I think we should, we are left with a problem that affects the entire sphere of human cognition—a problem which, therefore, cannot simply be overlooked in favour of different philosophical approaches. There may well be different kinds of truth, each valid in its own way, but if all the words, concepts and meanings employed by ordinary consciousness are potentially unknown and unintelligible, every particular truth, insofar as it rests on these very words, concepts and meanings, is liable to be compromised. And, if so, every form of freedom, insofar as it depends on truth and cognition *at all*, is likewise open to attack. If Plato is right, his criticism is not simply one among others, a philosophical angle to be adopted or disregarded. It is a truly foundational problem, challenging the very possibility of knowledge in all its modes and affiliations.

I will return to this issue in due time. For now, having clarified this book's main aims and methods, I would like to say something about its main sources. To my knowledge, the kind of analysis offered here has not been previously attempted. Still, other works have focused on different aspects of the debate under consideration, and many have helped shape the ideas and theses expounded below. First, with regard to Plato, there is a vast bibliography dedicated to his views on both knowledge and freedom. The relationship between these two themes, however, has not received such widespread attention. This is mainly because most studies on Plato's conception of freedom are centred on its political dimension, and particularly on his conception of the political state. A notable exception is Robert Muller's *La Doctrine*

Platonicienne de la Liberté, which offers a comprehensive reading of Plato's notion of freedom, retracing its various forms throughout the dialogues. Muller shows that freedom is an overarching Platonic theme, whose essence leads back to Plato's epistemology. Another exception is Carmelo Librizzi's *Platone o la Ricerca della Libertà*, which also treats the issue of freedom within the wider context of Plato's epistemology. Librizzi credits Plato with having been the first to grasp the fundamental importance of the problem of freedom 'with regard to the solution of all other philosophical problems.'⁷³ Muller's and Librizzi's approaches anticipate, in this respect, my own characterization of Plato's philosophical project.

The first part of this book is also indebted to the heterodox reading of Plato by authors like Mário Jorge de Carvalho, Francisco Gonzalez, Drew Hyland and, to some extent, Wolfgang Wieland.⁷⁴ For a long time, mainstream scholarship has reduced Plato's thought to 'Platonism', a 'grandiose metaphysical theory of "separately existing" Forms'⁷⁵ located somewhere beyond human experience. Having bound Plato to the rigid opposition between the 'real world' and a transcendent, heavenly realm, most interpreters lost sight of the *immanent* tension highlighted in the dialogues. And this binary reading, as one commentator rightly noted, 'has stood in the way of a real confrontation with philosophy as Plato understood and depicted it in his singularly elusive dialogues.'⁷⁶ Departing from this line of interpretation, the aforementioned authors have returned to the dialogues' original text to uncover a richer and much subtler philosophical project. The present work follows in their footsteps: instead of the usual focus on Plato's alleged dualism, his so-called 'two-world theory' and similar exegetical labels, I present his work as an open philosophical enterprise, aimed at questioning the beliefs and concepts we usually take for granted. I argue, furthermore, that only by releasing Plato from the preconceptions usually associated with 'Platonism' can his dialogues emerge as the foundation of a complex and fruitful philosophical tradition—the tradition examined throughout this book, leading all the way to Hegel and Nietzsche.

Second, regarding Plato's influence on these two philosophers, the existing literature is extremely varied. In Nietzsche's case, the link with Plato is obvious and well documented. Nietzsche himself saw his entire work as a 'struggle against Plato'⁷⁷ and directed many of his arguments specifically against Socrates, Plato and the Platonic tradition. This link was highlighted and discussed in many of the works I consulted while preparing this book—both in the general monographs by Jean Granier, Alexander Nehamas or Maudemarie Clark, among others, and in works specifically focused on the Nietzsche-Plato/Socrates debate, namely by Werner Dannhauser, Michel Guérin, Annamaria Lossi or Monique Dixsaut. In Hegel's

⁷³ Librizzi, *Platone o la Ricerca della Libertà*, 103

⁷⁴ References for these and the following authors can be found in the bibliography.

⁷⁵ Vlastos, *Socrates*, 48

⁷⁶ Gonzalez, *Dialectic and Dialogue*, 16

⁷⁷ BGE, 4 / KSA 5, 12

case, the connection with Plato has also given rise to a vast literature, including Louis Vieillard-Baron's classic study *Platon et l'Idéalisme Allemand* and Friedrich Kümmel's *Platon und Hegel*.⁷⁸ However, unlike the Plato-Nietzsche connection, Hegel's Platonic heritage has not been as widely explored. Although both authors are usually labelled as idealists and compared with regard to their use of dialectics, Hegel's (post-)Platonic approach to the relationship between freedom and knowledge has not been specifically addressed. And this gap is due, in large part, to the standard reading of Plato I have criticized above. Indeed, while Hegel's debt to Plato's dialectics is widely acknowledged, the emphasis is usually placed on the novelty of the former's approach. Whereas Plato is portrayed as a somewhat naïve metaphysical dualist, Hegel is hailed as an altogether different kind of thinker, bent on overcoming dualism and breaking away from the metaphysical tradition. In the present work, I challenge this attitude by offering new readings of both Plato's and Hegel's philosophical projects. I argue not only that Hegel's thought is indeed greatly indebted to Plato's, but also that the *Phenomenology of Spirit* can be read as a response to the main challenge raised in Plato's dialogues.

To this end, I will focus on a crucial point of contact between Plato's and Hegel's philosophies, namely the structural and thematic similarities between the *Republic*'s allegory of the cave and the *Phenomenology*'s dialectical progression. Even though some attention has been paid to this issue, its implications remain largely unexplored.⁷⁹ I will address this gap by highlighting the proximity of the two texts and their common aim of re-educating and liberating human cognition. Moreover, I will show that this double aim is also shared by Nietzsche, whose work includes important references to the allegory of the cave and the ascensional model developed by Plato and Hegel.⁸⁰ In all three authors, the image of the cave, as well as the contrast between slavery and freedom—along with a series of related images, such as the opposition between light and darkness, depth and altitude, dreaming and wakefulness, etc.—speak to an overarching concern with the cognitive foundations of human freedom.

Finally, the relationship between Hegel and Nietzsche has been the object of several studies and monographs. Despite the variety of readings put forward in the last decades, it has become a habit to divide them into two factions: the first one is led by Gilles Deleuze, whose influential, if questionable take on the Hegel-Nietzsche debate sees both philosophies as radically opposed to each other and fundamentally incompatible.⁸¹ This approach is shared, in different ways, by Derrida, Foucault or Habermas, and endorsed by most readers. The second faction, led by Walter

⁷⁸For a comprehensive bibliography on Hegel's relationship with Plato and Aristotle (up to the 1980s) see the list compiled by Norbert Waszek in Riedel, *Hegel und die antike Dialektik*, 275–283.

⁷⁹See fn. 46.

⁸⁰See Chap. 10, and Blumenberg, *Höhlenausgänge*, 613–636.

⁸¹Deleuze sees Nietzsche's entire philosophy as a reaction to Hegel and the Hegelian tradition. See Deleuze, *Nietzsche et la Philosophie*, 9, 187.

Kaufmann, advocates what Daniel Breazeale termed a ‘rapprochement’ interpretation of the relationship between both authors, centred not on their differences but on their similarities.⁸² My own analysis falls between these two approaches. Like Stephen Houlgate, Will Dudley or Robert Zimmerman, I believe Hegel’s and Nietzsche’s philosophies share significant thematic similarities, which warrant a comparative study of their main arguments.⁸³ Nevertheless, there is a crucial sense in which their philosophical strategies are indeed opposed to each other, and ultimately incompatible. To my mind, this ambivalence is best understood in light of Plato’s philosophy: although Hegel and Nietzsche agree with the Platonic indictment of natural reason and the need to re-educate and liberate human consciousness, each devised a different way of achieving these goals. And their solutions are not merely different: each ‘represents a decisive culminating and turning point in the history of Western thought and culture.’⁸⁴

Given the thematic proximity of a few earlier contributions to my own, some works merit a special mention. Among them, Will Dudley’s monograph *Hegel, Nietzsche, and Philosophy: Thinking Freedom* addresses some of the main issues discussed in this book. But while Dudley grounds his comparison of Hegel’s and Nietzsche’s views on freedom in their rejection of Kantian morality, I adopt a wider historical framework and interpret the Hegel-Nietzsche debate in light of the Platonic tradition, broadly construed. Moreover, whereas Dudley’s treatment of Hegel is centred on the Philosophy of Spirit (as expounded both in the *Philosophy of Right* and the *Encyclopaedia*) and the Logic, my main focus is the *Phenomenology of Spirit*. My aim is to go back to the beginning, as it were, and document the transition from consciousness’ natural state of imprisonment to a free and educated outlook on reality. Unlike the higher stages of Hegel’s system, the *Phenomenology* captures human freedom *in statu nascendi*. In Dudley’s own words, it provides ‘an account of the elevation of consciousness to a standpoint that Hegel characterizes as free’ and ‘prepares the way for the free conceptual thinking that transpires in the logic, the philosophy of nature, and the philosophy of spirit.’⁸⁵

In Stephen Houlgate’s *Hegel, Nietzsche and the Criticism of Metaphysics*, both philosophers are singled out as key turning points in the history of philosophy and ‘liberating alternatives to metaphysics.’⁸⁶ Houlgate compares Hegel and Nietzsche in light of their efforts to overcome the dualistic framework of classical philosophy, typically translated into the contrast between reason and reality, or a world of appearances and a world in itself. In this regard, his approach anticipates mine: in what follows I also present Hegel’s and Nietzsche’s philosophies as alternative

⁸²For a good survey of the existing literature on this topic (up to 1985), see Houlgate, *Hegel, Nietzsche and the Criticism of Metaphysics*, 1–23.

⁸³See also the articles by B. F. Beerling, D. Breazeale and F. Cauchi listed in the bibliography.

⁸⁴Breazeale, ‘The Hegel-Nietzsche Problem’, 147

⁸⁵Dudley, *Hegel, Nietzsche and Philosophy*, 15f. I discuss the transition from the *Phenomenology* to Hegel’s later works in Sect. 9.2.

⁸⁶Houlgate, *Hegel, Nietzsche and the Criticism of Metaphysics*, 2

responses to a fundamental cognitive tension, consistently thematized throughout the history of Western thought. Unlike Houlgate, however, I explicitly trace the source of this debate back to Plato and the Platonic tradition. I show that Plato's purported dualism can be read as the inaugural chapter of a long philosophical debate, which Hegel and Nietzsche will both pick up on and attempt to solve. Moreover, as already mentioned, I argue that both authors are further united by the realization that this debate is not merely logical or epistemological, but above all a debate about freedom and the possibility of an independent and meaningful life.

Particularly important in this context is the research devoted to Hegel's and Nietzsche's accounts of the master-slave relationship, and to their diverging assessments of the Judaeo-Christian tradition. In a comparative analysis of Hegel's 'unhappy consciousness' and Nietzsche's 'slave morality', Murray Greene highlighted the conceptual limitations of the latter's 'life-affirming' philosophy: by focusing on the individual affirmation of a 'noble' or 'masterly' consciousness, Nietzsche confined his criticism to a simple opposition, where both sides exclude one another. For Hegel, on the contrary, the conflict between the self and its other can only be resolved through a reciprocal exchange, grounded in a mutual act of recognition. Nietzsche's account lacks the moment of mediation that for Hegel is the key to overcoming the master-slave conflict, as well as the self-alienation of religion and the contradictions of scepticism, stoicism and modern reason.⁸⁷

More recently, Robert Williams revisited this issue by confronting Hegel and Nietzsche with Deleuze, to whom the Nietzschean notion of 'slave morality' represents a direct indictment of Hegel's master-slave dialectic.⁸⁸ For Deleuze, 'the overman is directed against the dialectical conception of humankind, and transvaluation against the dialectic of appropriation, or of suppressing alienation.'⁸⁹ Following Greene's criticism, Williams exposes the inconsistencies of Deleuze's Nietzschean critique of Hegel and highlights the latter's focus on the social foundation of human freedom. Whereas for Deleuze (and Nietzsche) the need for recognition is reactive and servile, characteristic of a weak and powerless consciousness, 'recognition for Hegel is that act and process through which the social itself is constituted, whether in unequal shapes of domination, e.g. master/slave, or in shapes of mutuality, e.g. as in reciprocal recognition and ethical life. The social is not essentially heteronomous, but ultimately a condition of freedom's actualization in the world.'⁹⁰

Although I agree with Greene's and Williams' criticisms, my aim is not merely to compare Hegel's and Nietzsche's views on power and recognition. My main goal is to determine how these views reflect both philosophers' wider strategy of educating and liberating human consciousness. Hegel's master-slave dialectic and Nietzsche's philosophy of power can be read, in my view, as different attempts to overcome

⁸⁷ See Greene, 'Hegel's "Unhappy Consciousness" and "Nietzsche's Slave Morality"'.

⁸⁸ Williams, *Tragedy, Recognition, and the Death of God*, 33–53

⁸⁹ Deleuze, *Nietzsche et la Philosophie*, 9

⁹⁰ Williams, *Tragedy, Recognition, and the Death of God*, 46

ordinary consciousness' usual state of bondage, rooted in its problematic relationship with truth. But while Nietzsche locates this release beyond the cognitive realm, in the direct experience of power and domination, Hegel's power struggle is but an intermediate stage in a longer emancipation process, whose ultimate goal is a true, and hence truly free outlook on reality. Again, I will show that this debate is not exclusively Hegelian or Nietzschean and that its origin can be traced back to Plato's dialogues. I will focus, in particular, on the *Gorgias*, where the dispute between Socrates and Callicles anticipates, in many respects, the rift between Hegel's *Geistphilosophie* and Nietzsche's *Machtphilosophie*. Like these two philosophical projects, Plato's dialogue is centred on human reason's natural impotence and on the need for a major cognitive transformation. But whereas Callicles' idea of freedom entails a release from knowledge and truth, close to Nietzsche's 'overhuman' emancipation, Socrates points out the contradictions of this kind of solution, paving the way for Hegel's own critique of mastery and domination.

In the end, however, my analysis will show that the problem raised by Plato is left largely unsolved. In Nietzsche's case, all things considered, the search for a post-cognitive brand of philosophy seems doomed to lead back to the cognitive basis he seeks to avoid. Just as Callicles' position is shown to be untenable, so too is Nietzsche's notion of power ultimately tied to some form of rational normativity. In Hegel's case, the solution to Plato's predicament is more sophisticated, but no less problematic. As he himself makes clear, the success of his approach hinges on an ambitious set of requirements, regarding the comprehensiveness, systematicity and necessity binding together the *Phenomenology's* different stages. Yet these requirements are open to different forms of contingency, liable to compromise the possibility of a dialectical liberation of human consciousness.

Part I
Plato's Philosophical Project

Chapter 2

Plato's Idea of Truth



Abstract Human freedom is often presented by Plato as a paradox. Although most men and women regard themselves as free and independent beings, this vision is open to refutation. Human life may in fact amount to a peculiar form of imprisonment and human beings to powerless creatures, unaware of their true condition. This idea, famously illustrated by the allegory of the cave, is also conveyed in other Platonic texts, and its scope is much wider than is generally assumed. The lack of freedom highlighted by Plato does not refer merely to the inability to achieve certain goals or to overcome certain obstacles. It stems from the lack of a true and clear account of reality, in which to ground one's judgements and decisions.

In order to clarify Plato's critique, this chapter summarizes his views on the structure and scope of human cognition, scattered throughout the dialogues. This initial characterization is meant to establish the basic meaning of key Platonic distinctions such as belief (δόξα) and science (ἐπιστήμη) or body (σῶμα) and soul (ψυχή), which will come up at different junctures throughout this book.

Keywords Plato · Socrates · Freedom · Truth · Doxa · Body · Soul

*Die Wahrheit!
Schwärmerischer Wahn eines Gottes!
Was geht die Menschen die Wahrheit an!*¹

In Plato's dialogues, human freedom is often presented in the form of a paradox. Although human life is usually assumed to be free, and although human beings regard themselves as independent and powerful beings, these beliefs are open to refutation. As Socrates keeps pointing out, things might just be the opposite of what they seem: human life may in fact be an imprisoned life and human beings powerless or enslaved beings, unaware of their true condition.

¹ KSA 1, 759: 'The truth! Rapturous illusion of a god! What does the truth matter to most human beings!'

This idea is famously illustrated in the *Republic*'s seventh book, in the so-called allegory of the cave. But this well-known image is not an isolated case. In other texts, Plato argues, for example, that human beings usually regard reality 'as through prison bars',² or that we are usually 'in the condition of common slaves.'³ Imprisonment, servitude and impotence play a central role in Plato's characterization of human life and their scope is much wider than most readers tend to acknowledge. The lack of freedom highlighted in the dialogues does not refer simply to the incapacity of achieving certain goals or overcoming certain obstacles or difficulties. It refers to the way we cognize reality and to the meaning we ascribe to objects and events. More specifically, it stems from the usual lack of a true and intelligible account of reality, in which to ground our judgments, choices and decisions.

To determine the exact implications of Plato's argument, let us start by considering his main views on the structure and scope of human cognition, scattered throughout his writings.

2.1 Truth, Opinion and Doxa

Plato's enquiries are usually focused on the lack of coincidence between what most of us think we know about reality and what we actually know about it. Their aim, however, is not simply to point out the finite or limited nature of human cognition. To be sure, all men and women are ignorant about a great number of things and uncertain or confused about a great many more. Yet the primary focus of Plato's critique is not this generic form of ignorance. His main concern is rather to show that human beings are ignorant and deluded about the very things they claim to know best.

Unlike other men, Plato's Socrates 'recognizes that he is ... of no account in respect to wisdom.'⁴ Convinced of his own ignorance, he sets out to put other people's knowledge to the test, and his painstaking enquiries invariably lead to the same conclusion: most people do not know what they claim to know; human wisdom, even as regards the simplest or most familiar notions, 'is of little or no value'.⁵

For Socrates, as for Plato, human consciousness' usual incomprehension of the most basic aspects of reality is primarily due to a basic incomprehension of itself. Though all human beings assume they know a great deal of things, and live their whole lives under that assumption, they are usually ignorant about what knowing

² *Phd* 82e2–4

³ *Sy* 215e6–7

⁴ *Ap* 23b3–4

⁵ *Ap* 23a7

actually means. Focused as they are on specific objects and events, they rarely take the time to look at how these things come to be cognized. The world's meaning and intelligibility are taken for granted and reason's efforts are usually saved for more concrete or complex problems. In Plato's dialogues, on the contrary, Socrates goes back to a more elementary form of enquiry. He turns his attention to the backstage of the cognitive process, in search of its underlying structure.

In everyday life, objects and events are perceived as facts to be witnessed and taken note of. Their identity, however vague, is usually thought of as something fixed and determined, and the circumstance of their being known does not alter in any significant way what they already are. Against this attitude, Plato shows that cognition is never simply about the direct contact with some original core of reality, but about the manipulation of a basic supply of ideas or concepts, through which the world is rendered intelligible and meaningful. Moreover, these concepts are never independent or isolated. As we have seen earlier, they are always part of a complex cognitive web, which can be organized in different ways and convey a potentially infinite array of meanings. Just as the identity of each object or event is the result of a complex combination of concepts, so too is the latter part of a wider cognitive web, where different concepts take up different hierarchical positions and assume different degrees of importance.

As Plato unearths this complex cognitive system, he calls attention to its global arbitrariness. And his critique is directed both at the system's alleged *validity* and at its alleged *intelligibility*. In what regards the first of these aspects, he points out that human reason is entirely dependent on the idea of truth. Indeed, it could be the case that we were somehow indifferent to the truth-value of our ideas and went about our business not caring whether things are in fact what we perceive them to be. But this is never the case. Our thoughts and judgements are always accompanied by an expectation of truthfulness: we expect our usual account of reality to be true, rather than false, and to reveal not incomplete or distorted images of things, but things as they really are.

This does not mean, of course, that we wonder ceaselessly about the truth-value of all our thoughts, nor that we are seized by endless philosophical doubts when performing the most basic everyday tasks. This is not the case because our usual account of reality is not interrogative, but assertive: given any object or event, our representation of it is never reduced to the representation of a mere object or a mere event; it also includes, simultaneously, the representation of its truth-value. In other words, whenever we perceive an object or an event, we perceive it to be *so* or *so*—*and not otherwise*. The conceptual web underlying each of our representations is automatically accompanied by a thesis regarding its own validity, without which objects and events would amount to pure possibilities, and not to actual things or states of affairs.

Furthermore, the cognitive web underlying our usual account of reality is also accompanied by a thesis regarding its own intelligibility, whose origin is no less arbitrary. Objects and events are usually familiar to us because their meaning rests on familiar concepts, combined in different ways. Again, however, just as these

concepts are merely asserted to be true, they are also merely assumed to be clear or intelligible. Once their meaning is questioned, what emerges is a peculiar form of indetermination. Since each different concept, or combination of concepts, is defined by its position within a global conceptual hierarchy, where all concepts are likewise defined by other concepts, every meaning lies outside of itself, as it were, and cannot be fully grasped.

In the dialogues, this double critique is usually conducted in one of two different ways. In some cases, Socrates' charges are triggered by the consideration of a specific concept, taken up at random or at the request of his interlocutors. The enquiry moves from a particular case to an increasingly wider conceptual debate, where reason's inner contradictions are gradually unveiled. In other occasions, Socrates' analysis is directly grounded in a global description of human cognition, where the same contradictions are discussed in a more systematic way. For practical reasons, I will ground my own exposition of Plato's argument in the second of these approaches.

With the previous considerations, we have arrived at what is arguably the central theme of the dialogues: the recurrent claim, phrased and illustrated in multiple ways, that reason is not a mere *taking note* of reality, as though the latter were immediately accessible or self-evident; cognition consists rather in a series of *theses* about reality, automatically assumed to be true. In Plato's vocabulary, these theses, grounded in more or less complex constellations of concepts, are called δόξαι, and their exact definition is crucial to understand the essence and scope of Plato's philosophical project.

First of all, to grasp what Plato means by δόξα and δοξάζειν we must avoid the false friends associated with the translation of these terms as 'opinion' (and 'to have an opinion' about something), 'view' (and 'to hold a view' about something), 'judgement' (and 'to judge' something to be so or so), etc. Through these and similar formulas, we risk contaminating Plato's notion of δόξα with our own modern notions of opinion or judgement, and thereby missing his entire point. This contamination stems above all from three main properties we ascribe to opinions or judgments, but which are absent from Plato's definition: namely, what might be called a) the *non-binding* nature of opinions; b) the *episodic* nature of opinions; and c) the *intelligible* nature of opinions.

The first of these properties refers to the degree of certainty enjoyed by opinions. Usually, when one holds a specific view about something, one is claiming that, as regards the object or state of affairs in question, *A* or *B* is the case—and not *C* or *D*. An opinion, however uncertain, is by definition the determination of a specific version of things. But precisely because it is an opinion, and nothing more than an opinion, it is unable to exclude the possibility that things might be otherwise. Unlike factual or apodictic truths, opinions convey provisional information, whose full confirmation requires further evidence.

Depending on the context, Plato resorts to what might be called a common use of the term δοξάζειν (and all related expressions) and a technical use of this term (and all related expressions). The first one, frequently found in many ancient authors, can indeed be translated as 'having a determinate opinion' or 'holding a determinate

view' about something. In this case, δόξαι amount to beliefs about reality whose degree of certainty can vary considerably. In the second sense, though—the Platonic sense *par excellence*—the term δοξάζειν refers to a much stronger kind of opinion. More precisely, to an opinion that is not seen or lived as an opinion, but as *the actual truth*. In this second case, δόξαι amount to beliefs exposed to the possibility of error, but which have nonetheless been raised to the status of factual or apodictic truths.⁶

The second way in which Plato's δόξαι differ from simple opinions concerns their scope. For most of us, opinions are frequent, but isolated events. They arise in connection with a specific object or state of affairs, amid a wider universe of objects which elicit no opinion on our part, either because we already know what they are or because we are indifferent to them. Either way, opinions are usually construed as explicit cognitive events, aimed at those segments of reality that interest or puzzle us, of which we have no definite knowledge. Once again, however, what is highlighted in the Platonic notions of δόξα and δοξάζειν is precisely the opposite, namely, the fact that human opinions are not episodic or isolated, but *pervasive*. By insisting on these notions, Plato is pointing out that human knowledge *is from top to bottom a matter of δοξάζειν*. Whenever we think we know something, we really only have an opinion about it—and one whose appearance of truth belies a disconcerting lack of evidence.

Third, Plato's δόξαι also differ from simple opinions in that they are ultimately unintelligible. When most people state an opinion they are usually convinced they know what they are talking about. The content of opinions may be uncertain or controversial, but the terms in which they are expressed are not. The opinions themselves, regardless of their truth-value, take the form of intelligible sentences, whose meaning is generally familiar to those who state them and to those who listen. Yet by calling attention to the *doxastic* nature of human cognition, Plato is highlighting the potential unintelligibility of human beliefs, and of the cognitive alphabet in which they are grounded.

To appreciate how these three differences come into play in the Platonic conception of δοξάζειν, let us turn to an important passage from the *Philebus*, where Socrates and Protarchus discuss the origin of human opinions.⁷ In order to explain how a given δόξα comes into being, Socrates resorts to the image of a man who sees a distant object and tries to make out its identity.

SOCRATES: Would you say that often when a man sees things at a distance and not very clearly, he wishes to distinguish between the things which he sees?

PROTARCHUS: Yes, I should say so.

SOCRATES: Next, then, would he not ask himself—

⁶In this second sense, the words δόξα and δοξάζειν may be used either in a neutral or in a pejorative way. In the first case, they emphasize a given judgement's categorical nature (as opposed to διστάζειν, which conveys the idea of doubt or indecision), regardless of its actual truth-value. In the second case, they apply to malformed or unreliable judgements, wrongly perceived as sound or solid. Nevertheless, as we will see shortly, this distinction is ultimately irrelevant, for every δόξα, according to Plato, is by definition a malformed judgement.

⁷See *Phi* 38c–39c.

PROTARCHUS: What?

SOCRATES: 'What is that which is visible standing beside the rock under a tree?' Do you not think a man might ask himself such a question if he saw such objects presented to his view?

PROTARCHUS: To be sure.

SOCRATES: And after that our gazer might reply to himself correctly 'It is a man'?

PROTARCHUS: Certainly.

SOCRATES: Or, again, perhaps he might be misled into the belief that it was a work of some shepherds, and then he would call the thing which he saw an image.

PROTARCHUS: Yes, indeed.⁸

This passage calls attention to a procedure that is usually invisible or automatic. Whereas most of our beliefs about reality take shape instantaneously, with little or no hesitation, the observer imagined by Socrates is purposefully held in a state of indecision, tempted by a series of compelling alternatives. At stake is not the end result of reason's representational process, but the process itself. What is brought to the fore is the very act of *δοξάζειν*, prior to the closing in on a determinate *δοξαζόμενον*.

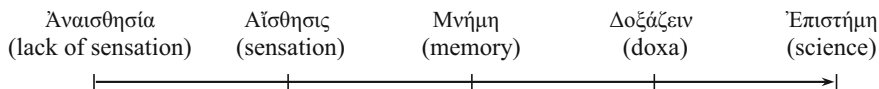
In Plato's words, every human belief is born from a *βούλεσθαι κρινεῖν*, that is, from the urge to distinguish between different accounts of what something is or means. Moreover, in the lines leading up to this passage, Socrates argues that this urge is triggered *ἐκ μνήμης τε καὶ αἰσθήσεως*, i.e. by the combined action of sense perception (*αἴσθησις*) and memory (*μνήμη*), through which reality is made available and cognizable in the first place.⁹ But this new claim must be read with care. At first sight, it might be taken to mean that sense perception and memory are the original sources of human cognition, which would contradict my previous insistence on the *pervasive* nature of *δοξάζειν*. Indeed, if there is something prior to the doxastic plane—a more primitive mode of representation, in which *δοξάζειν* itself is grounded—it is only natural to assume that cognition is *not* from top to bottom a matter of *δοξάζειν*. And, moreover, that it is not *primarily* a matter of *δοξάζειν*. If human opinions are indeed born from sense perception and stored by memory, the stage of opinion is but a secondary (or tertiary) stage, added to an already formed nucleus of sensory and mnemonic data. And, if so, it is also natural to assume that *δοξάζειν*, albeit responsible for validating and bringing into relation the correlates of sensation and memory (the various *αἰσθήματα* and *μνήμαι*), possesses but a limited range of variation. For surely sensations and memories, insofar as they are not originally doxastic, convey a kind of truth that is not as open to being questioned as the wavering truth of opinions. And surely the brightness of a determinate colour, for example, or the pitch of a determinate sound, or the smell of a determinate flower are not matters of opinion in the same way that the recognition of an object seen from a distance might be. Given the priority of sensation and memory in human consciousness' cognitive hierarchy, there seems to be an *immediate* or *factual* core of information that is independent from *δοξάζειν* and unaffected by it.

⁸ *Phi* 38c5–d11. See also *Phi* 42a, *So* 234b, *Th* 191b or *Pr* 356c.

⁹ *Phi* 38b12

Once more, however, this is precisely the kind of misconception Plato aims to expose. If men and women were not endowed with the ability to see, or hear, or touch, nor with the capacity to remember what they saw, heard, or touched, they would be unable to cognize reality. But Plato's point is that everything one is able to represent as *being something*, or as *having a determinate meaning*—and *a fortiori* one's representation of sense perception and memory themselves—is represented *from the viewpoint of* $\delta\omicron\zeta\acute{\alpha}\zeta\epsilon\upsilon\nu$. Although memory and sensation are presumed to underlie every human belief, their role can only be recognized and appreciated through the all-embracing filter of $\delta\omicron\zeta\acute{\alpha}\zeta\epsilon\upsilon\nu$. And this is so because human beings' usual standpoint is not a sensuous or mnemonic standpoint, but a *doxastic* standpoint.

In the *Philebus*, Socrates places our usual standpoint within a cognitive interval ranging from a minimum of cognitive insight, corresponding to a complete absence of cognition ($\alpha\nu\alpha\iota\sigma\theta\eta\sigma\acute{\iota}\alpha$), to a maximum of cognitive insight, corresponding to an epistemic or scientific outlook on reality ($\epsilon\pi\iota\sigma\tau\acute{\eta}\mu\eta$).¹⁰ This interval can be represented as follows:



The first main aspect to consider is the qualitative discontinuity of this progression. Inasmuch as every stage stands for a different cognitive regime, the transition from each one to the next entails a radical change of perspective. If a given consciousness were situated at the stage of $\mu\eta\acute{\eta}\mu\eta$, for example, it would be incapable of conceiving either the cognitive perspective of a consciousness situated at the stage of $\alpha\acute{\iota}\sigma\theta\eta\sigma\iota\varsigma$ or the cognitive perspective of a consciousness situated at the stage of $\delta\omicron\zeta\acute{\alpha}\zeta\epsilon\upsilon\nu$. In the first case, a mnemonic standpoint would be incapable of envisaging a purely sensuous account of reality because memory is something entirely different from the mere succession of sensations. In order to remember something, one must be familiar with the idea of duration and capable of distinguishing between past, present and future events. These two features are utterly inaccessible to a purely sensuous standpoint, to which reality is reduced to infinitely brief and unrelated flashes of information. Likewise, a doxastic standpoint is incapable of envisaging what a purely mnemonic standpoint might amount to because to hold a view about something is altogether different from storing and combining different sensations. To claim that something is *this* or *that* (is *this way* or *that way*, has *this nature* or *that nature*, etc.), one must be able to represent the very being of things and distinguish between what *is* and what *is not*, between truth and error, between adequate and inadequate portrayals of things, etc. All of this is inconceivable to a purely mnemonic standpoint, to which memories are nothing more than combinations of sensuous units, utterly alien to the notions of truth and untruth.

¹⁰ *Phi* 33c–34c, but also 21b–d, and *Th* 163c–164b.

When confronted with Plato's cognitive scale, every reader is led to form a mental image of what each of the stages stands for. The seemingly boundless plasticity of our usual standpoint leads us to believe we can shed our cognitive skin and put on radically different cognitive costumes. But this belief is, of course, illusory. As Plato points out, we are anchored all the while to the stage of *δοξάζειν* and fated to represent reality according to its parameters. When we consider the inferior stages of *ἀναισθησία*, *αἴσθησις* and *μνήμη*, we are really only considering our own doxastic standpoint *minus* the ability to distinguish between truth and falsehood, the ability to retain previously felt sensations and the ability to feel as such. In like manner, when we consider the stage of *ἐπιστήμη*, we are really only considering our own doxastic standpoint *plus* the ability to vouch for the truthfulness of our usual account of reality. In all of these situations, our standpoint remains as doxastic as ever, albeit disguised as a senseless, sensuous, mnemonic or epistemic standpoint.

It might be stated, in Kantian terms, that sensations and memories are the *matter* of human thoughts, whereas beliefs, or *δόξαι*, are its *form*. Let us start by focusing on the first half of this statement. Sensations and memories are the matter of human thoughts because without them cognition as such would be impossible. Without *αἰσθήματα* and *μνήμαι* there would be no fuel for *δοξάζειν*, and hence nothing whatsoever to assert or deny. The mind would have nothing to work with, no flesh to put on its opinions and beliefs. Nevertheless, *αἰσθήματα* and *μνήμαι* are by definition incapable of producing a determinate thought-content, whatever it may be. In the *Theaetetus*, when discussing the qualitative difference between sensation and knowledge, Socrates points out that the sight of a determinate colour or the audition of a determinate sound are not, by themselves, knowable thought-contents.¹¹ In order for a colour or a sound to be thinkable, one must be able to acknowledge that both of them exist, that each is different from the other and the same as itself, that both together are two and each separately is one, etc. None of these ideas are conveyed by the sheer perception of colours or sounds. They can only be acquired at a higher cognitive stage, which presupposes the silent action of memory and sensation but has no direct access to either of them.

Which brings us to the second half of my previous statement, namely that human beliefs are the *form* of human cognition. Since sensations and memories are silent as to the being and truth-value of what they represent, they must be *made to speak* by the animating power of *δοξάζειν*. And sure enough, returning to the *Philebus*, the man who sees an object from a distance and tries to discern what it is starts by asking questions to himself. Plato's use of direct speech to convey the observer's inner doubts reiterates the idea that the knowledge derived from the senses can only become knowable as a discourse. Only when what appears to the observer takes the form a question, and afterwards of a declaration or thesis, can it emerge as a determinate thought-content. We find this same idea in the *Theaetetus*, where

¹¹ See *Th* 185a–e.

Socrates argues that ‘the soul, . . . when it thinks, is merely conversing with itself, asking itself questions and answering, affirming and denying.’ And so ‘when it has arrived at a decision, whether slowly or with a sudden bound, . . . we call that its opinion [δόξαν]; and so I define forming opinion [δοξάζειν] as talking and opinion as talk which has been held . . . in silence with oneself.’¹²

The reason why Plato is so attached to this idea is now becoming clearer. At first glance, it seems fairly obvious that to perceive something one needs to be conscious that one perceives something, whether or not this acknowledgement takes the form of an ‘inner conversation’. But the point is that this acknowledgement is not a mere translation or clarification of a previously held knowledge about reality. Rather, it is the very source of human cognition: the ‘big bang’ which sets knowledge into motion, the ‘ground zero’ on which every impression and idea are made to stand. In the *Philebus*’ image, the *formal* priority of beliefs over sensations and memories is made clear by the fact that the object seen from a distance never presents itself simply as a cluster of sensations and memories. The unknown shape is perceived from the very beginning as one of different contending alternatives. Accordingly, the observer imagined by Socrates is not required to take note of a sensuous impression, but to decide between different theses about reality. His aim is to gather enough information to be able to settle for a determinate version of things—be it a coloured shape, a man or a scarecrow.

All of these elements lead to the same central conclusion: it is within the realm of δοξάζειν that human knowledge is born; and it is due to δοξάζειν that there is a world to be seen and heard, and objects to be distinguished and known. But if human cognition is indeed entirely made up of theses about reality, all human judgements are exposed to the possibility of error. And herein lies the true motivation for Plato’s concerns. If everything is what it is simply because it is asserted to be so, every single truth is susceptible, at least in principle, of being proven wrong. A cognitive system based on such shaky foundations is far from reliable—for despite all our expectations, there may be no actual connection between our representation of reality and what it is meant to represent.

In light of this extreme possibility, the metaphorical reach of the *Philebus*’ image appears even greater. Plato is not suggesting simply that objects seen from a distance are hard to identify, but that *all of our representations of reality*, insofar as they are based on the same doxastic procedure, are affected by the same kind of difficulty. In other words, whenever we consider any object whatsoever, we are bound to see it *at a distance*. Or, to borrow Goethe’s formula, ‘everything we possess, we see as if from afar.’¹³

If the situation imagined by Socrates were intended not as an allegory of human cognition in general, but merely as an illustration of a specific kind of cognitive experience—namely, the kind where the information retrieved by the senses is not sufficient or clear enough to trigger the adoption of a determinate belief—the

¹² *Th* 190a2–6. See also *So* 263e–264b.

¹³ Goethe, *Sämtliche Werke*, vol. 6.1, 536

observer's hesitation would not be particularly meaningful. This kind of predicament is not infrequent in everyday life and can be easily overcome by an increase in the amount of information provided by the senses. In Plato's example, if the distant shape approaches the observer, or vice-versa, its visual appearance will become more detailed and easier to identify. Moreover, it seems absurd to think that someone looking at another person standing only two feet away, for example, would mistake him or her for a scarecrow. In this and similar cases, the exposure to the possibility of error seems much smaller. And if one does get things wrong, one's mistake cannot be very significant.

It would seem, then, that there must be at least a hierarchical division between those beliefs that are more exposed to the possibility of error and those that are less. Even if one accepts that *all* of them are ultimately grounded in a mere assumption of truth, the belief in the existence of God, or the foretelling of a future event, for example, are surely more uncertain than one's belief that the sky is blue, or that one has a body. In the *Phaedrus*, Plato mentions the formal difference between controversial or 'abstract' concepts and more straightforward or 'concrete' ones. He even draws a distinction between the concepts whose meaning we all agree upon, like 'iron' or 'silver', and those whose meaning leads us to 'part company, and disagree with each other and with ourselves' (ἀμφισβητοῦμεν ἀλλήλοις τε καὶ ἡμῖν αὐτοῖς),¹⁴ such as 'justice', 'beauty', 'pleasure', 'piety', etc. Furthermore, it is a well-known fact that many of Plato's dialogues are explicitly centred on the discussion of precisely these 'abstract' or controversial concepts. Nevertheless, an attentive reading of any of the dialogues is bound to reveal that the distinction between self-evident and controversial concepts is ultimately spurious. In the course of Socrates' enquiries, the explicit ἀμφισβήτησις surrounding the notions of justice, beauty, pleasure or piety is invariably revealed as the keyhole from whence to peek at a much wider and more alarming form of ἀμφισβήτησις.

Usually, Plato resorts to what might be called a zooming device: starting from openly controversial notions, he proceeds to zoom in on their basic components. His aim is to eliminate the controversial nature of these concepts by isolating simpler concepts, devoid of ἀμφισβήτησις, and to use them as the ground on which to 're-construct' the controversial concepts. However, his hopes of finding such a ground are habitually shattered. The concepts of justice, beauty and the like are shown to rest upon simpler concepts that are themselves controversial, and these are in turn supported by yet another set of controversial concepts. In his attempts to zoom in on the most basic components of human cognition, what Socrates finds is not an original core of self-evident concepts, but a seemingly endless succession of controversial ones, with no firm ground in sight.

By stressing the potential disputability of consciousness' entire cognitive vocabulary, Plato is not simply arguing that we are unable to vouch for the truthfulness of our ideas. He is also suggesting that most of us, when employing either 'abstract' concepts like 'justice' or 'concrete' concepts like 'chair', are in fact unable to grasp

¹⁴ *Phdr* 263a2–b2

their actual meaning. When we claim, for example, that ‘justice is the noblest of virtues’; or when we say, for example, that a ‘chair is green’, we usually assume that we know what we mean—viz. that although we do not know everything there is to know about justice or chairs, we know enough to speak about them in a clear and unproblematic way. Yet when forced to explain what justice is, or what a chair is, we immediately resort to a chain of interconnected theses whose ultimate meaning is as unclear as that of justice or chair. In the first case, we say, for instance, that justice is to ‘accord everyone his or her share’, and that one’s share is ‘that which best suits his or her needs’; and that the best is the ‘highest degree of good’; and that good is ‘the opposite of bad’, and so on. In the second case, we say that a chair is a ‘solid object’, and that solidity is the ‘degree of resistance to motion’, and that motion is the ‘change of location over time’, and that time is ‘a measure of physical change’, etc. In both examples, the definitions offered amount to mere δόξαι, which are in turn defined by more δόξαι. Their meaning is caught up in a self-enclosed and potentially unintelligible maze of δόξαι, with no clear exit in sight.

In view of this, the amount of beliefs that may lead us to ‘part company, and disagree with each other and with ourselves’ is liable to be much greater than we tend to acknowledge. Plato’s zooming device suggests that we are ultimately unable to reach a consensus not only regarding the meaning of openly controversial concepts, but also regarding the meaning of much more straightforward concepts such as ‘iron’, ‘silver’, ‘large’, ‘small’, ‘table’, ‘chair’, etc. And this inability brings us back to the perplexity caused by the *Philebus*’ allegory: while we have no difficulty in accepting that the former set of concepts is indeed controversial and open to interpretation, we are not as open-minded when it comes to the latter. This is so, Plato argues, because our usual standpoint is not used to zooming in on its own concepts. Its habitual perspective is distant and fuzzy, just like that of the *Philebus*’ observer. It deals with concepts without actually seeing them or pausing to examine what they actually mean.

In short, Plato’s image does not refer simply to distant shapes or objects, but also to the very things placed in front of us. To borrow further from Goethe, ‘the hardest thing of all is that which seems easiest: to see what lies before our eyes.’¹⁵ And the difficulty exposed in the dialogues can be described in similar terms. It does not depend, ultimately, on the kinds of objects we perceive, or on the concrete conditions in which we perceive them. It is not derived from mere ignorance, or from lack of experience, or from some other contingent factor added to an otherwise sound cognitive standpoint. What Plato keeps highlighting is a structural problem, stemming from the very nature of our usual mode of cognition and liable to affect all our ideas, judgements and decisions.

¹⁵ Goethe, *Sämtliche Werke*, vol. 4.1, 694

2.2 Self-Confidence and Self-Interest

Socrates' usual task in the dialogues is to expose the doxastic foundation of human cognition. Throughout his enquiries, what is usually perceived as natural and self-evident is shown to be neither: familiar ideas and concepts, when subjected to a closer examination, are often reduced to a string of contradictions. But the dialogues also document the resistance put up by most of us to this kind of criticism. If reality as we know it is grounded in a system of theses, it is certainly not lived or experienced as such. On the contrary, human life is continually supported by the conviction that reality is indeed what it appears to be. And this conviction is so natural and deep-rooted that any suggestion to the contrary is readily dismissed.

It might be argued, against this general diagnosis, that human beings are not usually as confident about their cognitive powers. In everyday life, there seems to be much more room for error and illusion than there is for certainty, and though we are indeed certain of a good number of things, we often admit to being wrong or deluded about many others. Furthermore, we are usually ready to concede that what we do know is very little when compared with everything there is to know. Our limited grasp of reality, while enabling the performance of most of the practical and intellectual tasks required of us, is regarded as a small cognitive island amid a vast ocean of ignorance.

At first glance, this description seems accurate. On closer inspection, however, things are not quite so simple. Even though our usual knowledge is perceived as a small island, what lies outside of it is never purely unknown. If that were the case, the identity of distant objects and realities would be thoroughly indeterminate, open to an infinite range of possibilities: they might exist or not, be identical to the objects and realities we are familiar with or radically different, etc. Yet these possibilities are never truly undecided. We may have no experience of what lies outside our immediate cognitive horizon, but we do have an *idea* of it. And this idea is much less vague than we tend to acknowledge.

In reality, what we do not know or have no access to is never regarded as a pure mystery. On the contrary, it is always perceived as an extension of what we have already experienced. Our usual standpoint is guided by a general presumption of *homogeneity*, that is, by the general assumption that all empirical contents, whatever they may be, are bound to the same fundamental principles and endowed with the same fundamental properties. When we think of a distant planet, for example, we base our idea of it in the rules and qualities that define our own planet. When we picture the inside of our body, we do so based on its outer appearance. The basic characteristics of the things we are familiar with are automatically projected into all other regions of reality and transformed into global phenomenal principles. Distant or inaccessible objects, regardless of their actual nature, are never thought of as being colourless, for instance, or shapeless, or weightless, or devoid of texture, etc. Whatever they might be, they are always represented as having a determinate colour, a determinate shape, a determinate weight and a determinate texture.

Simultaneously, however, our usual standpoint is also guided by a general presumption of *heterogeneity*, that is, by the general assumption that whatever lies beyond our immediate cognitive horizon is different from what is contained in it. And, moreover, by the assumption that the various things located beyond our empirical horizon are as different from each other as those we are familiar with. If we were walking through a village and suddenly realized that every street is exactly the same, we would surely be amazed. Likewise, if we were to find out that every single house, in the same street, is perfectly identical to the next, we would also be perplexed. We assume that all things share the same fundamental qualities, but we also expect them to be different and varied.¹⁶

Yet these two assumptions are not incompatible. The differences we expect to find beyond our empirical horizon are not radical differences, but limited ones. And this is so because our usual presumption of heterogeneity is subordinated to our usual presumption of homogeneity, and not the other way around. Whereas the latter refers to basic properties or categories (colour, shape, texture, etc.), the former refers to specific instantiations of these categories (different colours, shapes, textures, etc.). What we expect to find beyond our empirical horizon are different variations of the same categorical repertoire. And since these variations are inspired, once again, by those we usually find in the things we are familiar with, the principle of heterogeneity is really only another instantiation of the principle of homogeneity.

This brief description is enough to disprove our initial hypothesis. Instead of an island surrounded by a sea of ignorance, our usual knowledge resembles a wide continent, or a wide cognitive web, extending in every possible direction. What we tend to regard as an uncharted territory is in fact already peopled with all sorts of beliefs regarding every aspect of reality. Indeed, instead of projecting *some* of the principles and qualities that define our empirical horizon into distant or inaccessible regions of space, we usually do so with regard to *all of them*. And, moreover, instead of projecting these principles and qualities into *some* regions of space, leaving others unexplored, we do so into *all of them*. Every reality we can think of is covered with assumptions or beliefs (δόξαι)—highly complex bundles of beliefs, concerning the world's colour and shape, the laws of motion and causality, the unity and hardness of physical objects, etc. This does not mean, of course, that we have as clear a picture of what goes on before us as we do of what goes on in distant planets or inside our body. But it does mean that these realities hold a very limited degree of surprise: what we do not know about them is only how the different beliefs we ascribe to everything else are combined.

¹⁶Leibniz called these two principles the principle of uniformity (*le principe d'Arlequin*) and the principle of variation (*le principe du Tasse*). See, for ex., *Die philosophischen Schriften*, vol. 3, 348; vol. 6, 548. While the first principle states that everything we know or experience is fundamentally similar to everything that exists, the second one states that everything is nonetheless different from everything else. Nature is both uniform and varied because every particular difference amounts to a different expression of the same totality. But while for Leibniz these two principles refer to the universe as it really is, they are being used here to characterize human consciousness' usual *assumptions* about reality.

Furthermore, these considerations do not apply only to the element of space, but also to the element of time. Apart from projecting our understanding of nearby objects and realities into other regions of space, we also tend to project our past and present knowledge into the future. When we consider any object, we automatically assume it will go on existing after the present moment. If a book lying in front of us would suddenly vanish into thin air, we would not fail to be amazed. Moreover, when we consider any object, we also assume its colour will remain the same after the present moment, as will its shape, size and texture. If the same book were to suddenly change colour, or lose its usual rigidity, or shrink to half its size, we would not be any less surprised. Finally, just as these things are not expected to happen to the book in front of us, neither are they expected with regard to any other book, anywhere else in the world. Here, again, our doxastic web extends in every possible direction. It is globally prescriptive, setting the ground for every future insight or experience.

The same holds, in addition, for the representation of our own identity. Throughout our lives, no temporal moment is lived as an isolated moment, or as the last moment there is. Besides assuming that the world will go on existing, in the same moulds and with the same general appearance, we also assume *we* will go on existing and experiencing reality as we have always done. What is more, the temporal horizon we place in front of us is not finite, but infinite. Even though we know we will eventually die, this knowledge does not prevent us from assuming, at each given moment, that a new moment will come, and a new one after that. Our thoughts and judgements, albeit conditioned by the idea of death, are always supported by the expectation of an endless temporal horizon, wherein all the beliefs we have previously considered are maintained and multiplied.¹⁷

The previous remarks are already indicative of the extraordinary complexity of our usual cognitive standpoint. Once again, however, all of the above is still but a part of a more complicated picture. Apart from all of our presuppositions regarding the world's spatio-temporal structure, a second and equally important set of presuppositions is also at work within us. Although we are usually guided by countless assumptions about the appearance and duration of objects and events, it might be the case that all of this were merely *witnessed* by us, in a neutral or disinterested way. If that were so, every object and event would be equally important and equally worthy of our attention. But this, of course, is never the case. We are never mere spectators of reality, but active and self-interested participants, for whom there is always something at stake. Our life has the form of a task, or rather of a highly complex set of tasks, whose fulfilment hinges on vast series of beliefs, judgements and decisions. Accordingly, our usual priority is not the acquisition of an epistemically pure account of reality, but one that is good enough to guide our actions and satisfy our practical demands. The criteria by which we determine things to be so or so, to be true or false, to have this or that nature, are not primarily epistemic, but practical and self-interested.

¹⁷ As Hans Blumenberg pithily puts it, 'We *know* that we must die but we do not *believe* it, because we cannot *conceive* it.' (*Höhlenausgänge*, 11)

Shaped by this additional set of judgements, the world is divided into things that matter and things that do not, things that are worth knowing and doing and things that are not, things that are beneficial and things that are harmful, necessary and superfluous, pleasant and unpleasant, safe and dangerous, etc. And these distinctions shed a new and important light on the principles of homogeneity and heterogeneity mentioned above. Even though our doxastic web extends to every corner of the universe, our idea of what goes on in distant planets, for example, is usually not as detailed or as demanding as our idea of what goes on in our street, or in our house, or in our living room. And the reason for this is not only that our senses are too limited to reach such distant realities. This cognitive asymmetry is due, above all, to our lack of interest in them. More specifically, it is due to the assumption that what goes on in other planets has no direct influence on the tasks and projects we are focused on. Nothing of what exists *there* is liable to alter our immediate perception of what goes on *here*, or to solve the concrete problems we are faced with every day. Then again, all of this might change if our practical horizon were different: if we were astronomers, for example, and earned a living from studying other planets, our interest in them would probably increase; if we were forced to abandon the Earth, due to some natural catastrophe, and to move to another planet, our perception of it would also radically change.

Our usual account of reality, albeit universal in scope, is centred on a specific cognitive perimeter. It is here that our life plays out and that our daily tasks are decided and executed. Everything falling outside this perimeter is deemed irrelevant or already settled. It is part of the background against which our life unfolds, and treated accordingly. However, as suggested in the previous example, this perimeter is not fixed or final. What lies in the background can always be brought to the fore, and vice-versa. This cognitive balance is dependent on a highly complex constellation of assumptions, whose inner composition is permanently updated in response to new contexts and situations.

In Plato's dialogues, Socrates calls attention to both of the main features I have briefly described. He emphasizes both the doxastic foundation of our spatial and temporal representations and their practical or self-interested nature. But his characterization of human cognition also includes a third fundamental element, in which the other two are grounded. And this new element also contradicts our understanding of our usual cognitive perspective. Normally, we tend to conceive the different tasks envisaged and performed every day as separate events, oriented towards different purposes. But this is never truly the case. All of our beliefs, despite their extraordinary diversity, are ultimately directed towards a single end. Human life may be diverse and incoherent, torn between different and often conflicting aims; and these aims may change with time, with circumstances, sometimes with mere chance. But all of them share a commitment to what is perceived, at each given moment, as the best or most advantageous course of action. Everything is chosen in view of what is deemed best—either directly, by aiming straight at what one wants, or indirectly, by aiming at something else, judged the best vehicle to attain what is best. This, according, to Plato, is the real source of human beings' constitutive self-interest.

The doxastic web that supports human life is organized in the form of a global scale, or a global pyramid, whose highest term is 'the good' (τὸ ἀγαθόν) or 'the best' or 'most excellent' (τὸ μέγιστον ἀγαθόν, τὸ βέλτιστον) of all contents or determinations.¹⁸ This crowning element is that in view of which everything else is thought of, judged, envisaged, decided, attempted or pursued. The greatest good does not refer only to big decisions or life-changing events. All daily actions and resolutions, from the most insignificant to the most important, are grounded in a specific version of it. 'It is in pursuit of the good that we walk, when we walk, conceiving it to be better; or on the contrary, stand, when we stand, for the sake of the same thing, the good.'¹⁹ Τὸ μέγιστον ἀγαθόν is the ultimate telos of human existence, the driving force behind all forms of cognition and action. Furthermore, the effects of this force are not occasional or intermittent. We are not interested in what is best only for a while, or in a limited way. We long to enjoy it *for as long as possible and in the highest possible degree*.²⁰

But this is still only one half of Plato's argument. Apart from bringing out the absolute priority of τὸ ἀγαθόν over everything else, he also points out that 'the good' is by definition a *formal* notion, liable to be translated into a multitude of different contents. On the one hand, 'nothing is more certain . . . than that every intelligent being pursues [the good], desires it, wishes to catch and get possession of it, and has no interest in anything in which the good is not included.'²¹ On the other hand, since the good can stand for almost anything, the pursuit of the good can also assume very different forms. What drives human life, therefore, are different versions of τὸ ἀγαθόν, whose content depends on each person's circumstances, interests and aspirations. Or rather, what drives human life is a global doxastic web entirely subordinated to what is assumed by each person, at each given moment, to be the best possible life.

By insisting on the formal nature of the good, Plato is arguing that the crowning element of our doxastic edifice is itself a doxastic judgement, prone to the same contradictions that affect all other human beliefs. At the same time, he is pointing out that this problem is not usually apparent to us, since our definition of the good is also automatic and taken for granted. In everyday life, instead of wondering ceaselessly about what the greatest good is, or how to obtain it, we simply go on living, focusing on specific tasks and projects. This is not to say that we do not have doubts about what to do or how to act, nor that we never wonder about the meaning of our lives. But these occasions are not that frequent; and even when they do constitute real moments of indecision, the latter usually affect only a very small part of our cognitive horizon. We may have trouble placing a specific belief, or set of beliefs, in our global cognitive edifice, but all of its other elements are left untouched. All other δόξαι remain committed, in one way or another, to the same version of τὸ ἀγαθόν.

¹⁸ See Sect. 1.3, fn. 29.

¹⁹ *Go* 468b

²⁰ See *Sy* 206a.

²¹ *Phi* 20d8–10

Plato's point, however, is not simply that the good can stand for different things. By insisting on its formal nature, he is also underlining its perpetual or self-renewing nature. Precisely because the good as such is always *better* than anything else, no concrete content is good enough to exhaust its meaning. One may dedicate one's entire life to saving up a million dollars, or buying a specific car, or having a specific number of children, but all of these goals, once achieved, are immediately surpassed by the idea of something even better, still left to accomplish. Since life's ultimate goal is formal, rather than concrete, we are all doomed to this permanent dissatisfaction. Every particular experience propels us to the next, which is in turn a prelude to something else.

Having outlined the main elements of Plato's conception of cognition, let us now return to the *Philebus*' thought experiment, where the identification of a distant shape encapsulates the complex cognitive structure we have been considering. To begin with, the observer imagined by Socrates is not baffled by the absolute novelty of what he sees. The object whose meaning he tries to make out is surrounded by a multitude of familiar shapes, whose meaning has already been decided. The theses hazarded about the shape's identity (i.e. whether it is a man or a scarecrow built by shepherds) are already framed by a myriad of other related theses, concerning not only the observer himself, his surroundings and his past experience, but also the very structure of the objective world. Regardless of what the shape turns out to be, it is already perceived as a solid body, with a specific colour, taking up a specific amount of space, subject to the laws of gravity and motion, etc. The shape's entrance into the observer's horizon is not a pure φαίνεσθαι, but the addition of a new element to a preformed version of things.

This idea is made clearer in the dialogue's immediate continuation, as Socrates compares human cognition to the joint work of a scribe (γραμματεὺς) and a painter (ζωγράφος), both active within the human soul.²² After suggesting that the 'matter' of our representations are sensations and memories, and their 'form' beliefs or δόξαι, in the terms I have previously discussed, he goes on to compare the 'inner discourse' that accompanies the embracement of different δόξαι to the double task of 'writing words in our souls'²³ and painting therein 'the images of what was established and said.'²⁴ In both of these metaphors, consciousness is likened to a sort of book or document (βιβλίον) in which the various δόξαι are recorded and combined.

At this point, Socrates adds a new and important element to his description. He observes that human cognition is not concerned only with the past and the present, but also with the future, and hence that the words and images written and painted in our souls concern not only things that *are* or *were*, but also things that *will be*:

SOCRATES: Do the writings and pictures, then, which we imagined a little while ago to exist within us, relate to the past and present, but not to the future?

PROTARCHUS: To the future, especially.

²² *Phi* 38e–39e

²³ *Phi* 39a3. Writing words, sentences or whole discourses—the word is λόγοι.

²⁴ *Phi* 39c6–7

SOCRATES: Do you say 'to the future especially' because they are all hopes [ἐλπίδες] relating to the future and we are always filled with hopes all our lives?

PROTARCHUS: Precisely.²⁵

The vital book Socrates uses as metaphor is not a mere depository of facts, past and present, as they are witnessed and taken note of. The words and images recorded by consciousness refer to memories and direct observations, but also to future situations and events. In order to grasp Plato's point, it is important to note that the word ἐλπίς stands not only for hopes or prospects—that is, for the desire for things to happen or evolve in a certain way—but also for expectations, grounded in presuppositions—that is, for the assumption that things will remain or become *so* or *so*, and *not otherwise*. The ἐλπίδες written and painted in our cognitive book concern not only how things *should turn out*, but above all how things *are supposed to turn out*, given their past and present condition.

Returning to the *Philebus*' initial example, the vision of a distant shape would not be possible (or, at least, would have a radically different form) without a series of contextual presuppositions relating not only to past and present, but also to future events. As the observer struggles to make out what he sees, he assumes, for example, that he will go on living after the present moment, that the world around him will go on existing, that his immediate surroundings will retain their colour and shape, etc. All of these beliefs, along with all the others mentioned above, make up the setting in which the distant shape makes its appearance. They provide the cognitive text, or the cognitive painting, as it were, in which the new shape is to be included.

By comparing consciousness to a book and by emphasizing the complexity of its content, Plato is highlighting an aspect of human cognition that is usually misunderstood. In his image, the scribe and the painter are continually at work, recording every new idea, object and event. However, their registry does not consist in a mere collection of ideas, objects and events, taken down in the order in which they are experienced. If this were so, the book of consciousness would be reduced to a journal of human experience, starting at the beginning of one's conscious life and ending in one's death. Whereas the first chapters would be comparatively empty, in accordance with the empirical poverty of infancy and childhood, the last chapters would be full of words and images, reflecting the empirical wealth of a long and seasoned life.

What Plato seems to be describing, however, is something different. The word he uses is not βιβλος, i.e. a book or a roll of papyrus, but a diminutive form, βιβλίον, which evokes the image of a single paper or tablet. And this difference suggests a different form of registry. The scribe's task is not to write at length, recording one sentence after another, but to *rewrite*, again and again, the same original text. Likewise, the painter's task is not to paint different images, one after the other, but to *repaint*, again and again, the same original one. Unlike the words and the images of a regular book or journal, the words and images mentioned by Plato are not definitive. While some are maintained or repeated, confirmed by the course of experience, others are erased and corrected, to accommodate new versions of the same ideas.

²⁵ *Phi* 39d7–e7

This paper or tablet is best described, therefore, as a palimpsest, whose changing content reflects consciousness' permanent cognitive updates. By insisting on the idea of a *complete* cognitive text, or a *complete* cognitive painting, continually reiterated and redefined, Plato is emphasizing the *complete* or *global* nature of our usual account of reality. The βιβλίον of human consciousness is already, from the very start, a *vital map*, offering a complete, albeit provisional picture of everything there is, was and ever will be.

But the δόξαι written and painted in our cognitive tablet do not refer simply to objects and events. As seen earlier, they also document our aims and aspirations, as well as our overall assessment of what is best or most important in life. And this practical dimension can also be illustrated with the aid of the *Philebus*' thought experiment. In the beginning, the distant shape is only a blurred outline. As it moves closer, however, its interference in the observer's practical horizon changes, and so does the kind of definition it elicits. A series of aspects that were previously irrelevant, like its actual size, its exact colour or its ability to move or speak become increasingly noteworthy, leading to the emergence of new assumptions and to the recycling of old ones. Conversely, as the shape moves away, or ceases to intrigue the observer, the composition of his vital map is once again transformed. As his priorities change, his attention is caught by other, more pressing concerns, and the same cognitive totality is reorganized. New objects and events are brought to the fore, while the old ones are gradually demoted. The shape's definition becomes less detailed and less important, until it eventually fades into the background.

2.3 Body and Soul

Plato's characterization of human cognition is remarkably modern. The idea that one's knowledge is determined by a continual urge for practical orientation, grounded in a global appreciation of what is at stake in one's life, seems to lead as far as the phenomenological and the existential traditions. And even though Plato does not speak of self-interest and practical reason in the explicit way Kant, Husserl or Heidegger would later do, he addresses the same general issues by dint of a more intuitive language. In many of his dialogues, the distinction between the doxastic nature of human cognition and the possibility of a true or epistemic account of reality is often framed by another, equally important distinction: the contrast between σῶμα and ψυχή. Once again, however, to understand what Plato means by these words, it is important to dispel a few common misconceptions.

In the dialogues, as in most classical Greek texts, σῶμα and ψυχή are usually translated as 'body' and 'soul'. But although this translation is undoubtedly correct, it does not get us very far—not so much because the words 'body' and 'soul' are in themselves incapable of conveying the meaning of this conceptual opposition, but because they are contaminated by a series of misleading assumptions. For modern readers, σῶμα usually refers to the human body with its members, organs

and concrete shape, whereas *ψυχή* evokes some obscure and ineffable entity lodged somewhere inside it. While the body can be seen and touched, the soul deals in ideas and feelings. But body and soul, however different, are both regarded as *things*, i.e. as the basic ingredients of a collection. When thrown together, they somehow blend into one another, producing a mysterious construct called life.

For Plato, however, *σῶμα* and *ψυχή* are not different things, but different ways of looking at reality, or different forms of interpreting human life. To reduce these terms to the elements of a collection is the surest way to lose sight of their Platonic meaning, and this is so for two main reasons. First, this kind of description relies on the possibility of a neutral or disinterested observer, placed outside the opposition and capable of a clear and impartial understanding of both terms. Plato's point, though, as we will see below, is precisely that such an observer does not exist. Since all of us are naturally endowed with a soul and a body, we are doomed to see reality from within this opposition. Our understanding of the soul and the body is itself conditioned by an original mixture of 'somatic' (*σωματικά*) and 'psychic' (*ψυχικά*) concerns. We do not know what a soul really is or what a body really is because we lack the distance required for such an assessment. A pure body and a pure soul can only be known indirectly, as the two boundaries within which our usual standpoint operates.

Second, by construing *σῶμα* and *ψυχή* as the elements of a collection, we fail to realize just how different they really are. There can be no symmetry between *σῶμα* and *ψυχή* because their relationship is not one of coordination, but one of subordination—and, moreover, one of *absolute* subordination. Recall Socrates' recurrent claim that the soul rules the body, and not the other way around.²⁶ If this is taken to mean that the soul rules the body as a thing rules another thing, we are led back to our initial dualism. But if Plato's insistence on the qualitative difference between soul and body is taken seriously, what emerges is a more radical idea: instead of an obscure entity lodged inside the body, the soul is the very condition for the body to be thought of as such; nay, it is the source of meaning *in general*, the founding principle of human consciousness.²⁷

By emphasizing the ruling nature of *ψυχή*, Plato is calling attention to the soul's absolute cognitive priority. Or rather, borrowing a term coined long after his lifetime, he is calling attention to the soul's *transcendental* nature. In his own language, he is arguing that cognition is by definition a 'psychic' event, or that reality is revealed, primarily, *through the soul*, or *with the eyes of the soul* (*διὰ τῆς ψυχῆς*). Therefore, the soul's limits are not those of the body, or of any other finite thing. They coincide with the limits of reality and extend as far as cognition itself. In other words, *ψυχή* is not a part of reality, but the window from whence to look out on it.

²⁶See *Phd* 94b, *Phdr* 246b, *Re* 353d.

²⁷This point is clearly made by Hegel in LHP, vol. 2, 37 / HW 19, 47f.: 'We assume that the soul can be or subsist without having imagination, thought, etc.; and the soul's imperishability is thereby regarded as the imperishability of a thing, of something represented as a subsisting being. For Plato, on the contrary, . . . thinking is not a quality of the soul, but its substance. The soul *is* the very thinking itself.' (Translation modified.)

In addition, by insisting on the ruling nature of ψυχή, Plato is also saying something about the kind of cognitive access it provides. In its pure or essential form, ψυχή is not simply the ability to cognize reality, but to do so in a thoroughly clear and unhindered way. A purely ‘psychic’ standpoint amounts to something like a global and absolutely revealing glance on the whole of reality. But things are very different in the concrete element of human life, where cognition is never as clear or as revealing. Unlike a purely ‘psychic’ standpoint, our usual standpoint amounts to a permanent struggle. It is continually forced to decide whether reality is so or so, to distinguish between true and false versions of things, to interpret ambiguous meanings and to make sense of obscure objects and events. Furthermore, it is usually focused on very small portions of reality and forced to make up for its ignorance through an endless series of assumptions and oversimplifications.

According to Plato, our usual standpoint is never purely ‘psychic’ because our soul is always accompanied by a body. Each of us is a ‘body and a soul put together’²⁸ and our outlook on reality is determined by the specific balance of forces struck by these two components. Unlike the soul, however, the body is devoid of cognitive insight. Whereas the soul is the primary source of knowledge, the body is a mere burden, a dead weight. It is *that in spite of which* the soul attains a clear understanding of reality. And yet its sheer presence is enough to alter the soul’s outlook in profound and significant ways.

If the soul were to suffer no interference from the body, its vision would reveal the world as it really is, in all its breadth and meaning. When in contact with the body, the soul’s universal gaze is reduced to a regional and unreliable form of access. It is forced to cognize the world through a restricted set of bodily organs, whose account of reality is limited to a restricted cognitive horizon and bound to a specific kind of access (viz. the access enabled by sight, smell, hearing, touch or taste). Moreover, the soul has to attend to a series of concrete and pressing bodily needs, ranging from the need to breathe, for example, to the specific challenges of locomotion, hunger or sexuality. As a result, the soul’s contemplation of reality is replaced by a self-interested concern with the body—with its preservation and welfare, with its practical orientation, with its enjoyment and satisfaction. In contrast to the soul’s original purity, this embodied standpoint is restless and passionate. And instead of the soul’s original immobility, it is continually led from pleasure to pain and from pain to pleasure.

Since the soul is usually confined to the body, it must view the world *with the eyes of the body* (διὰ τοῦ σώματος) and give in to its demands. In the dialogues, this situation is often portrayed as a form of imprisonment or servitude. In the *Cratylus*, Socrates associates the noun σῶμα with the verb σώζειν (to ‘preserve’ or ‘keep safe’) and defines the physical body as ‘an enclosure to keep [the soul] safe, like a prison.’²⁹ In the *Phaedrus*, he claims that the soul is ‘imprisoned’ in the body ‘like an oyster in its shell.’³⁰ In the *Phaedo*, a pure soul is said to be ‘freed from the body

²⁸ *Phdr* 246c5

²⁹ *Cra* 400c6–7

³⁰ *Phdr* 250c5–6

as from fetters',³¹ and afterwards a regular soul is described as 'entirely fastened and welded to the body and . . . compelled to regard realities through the body as through prison bars, not with its own unhindered vision.'³² As a result, 'each pleasure or pain nails it as with a nail to the body and rivets it on and makes it corporeal, so that it fancies the things are true which the body says are true.'³³

Human life, insofar as it depends on the human body, is not a free life. And this lack of freedom is grounded, once again, in a peculiar cognitive impairment. Yet Plato's contrast between the soul's intrinsic clairvoyance and the body's intrinsic ignorance does not amount to a simple antithesis. Rather, he draws up an entire cognitive scale, based on the negative idea of hindrance or interference. More precisely, he places human life inside a cognitive interval whose upper limit corresponds to the soul's originally unhindered vision (viz. the complete independence of $\psi\upsilon\chi\eta$ from $\sigma\tilde{\omega}\mu\alpha$) and whose lower limit corresponds to an absolutely hindered, and therefore utterly blind outlook on reality (viz. the complete subordination of $\psi\upsilon\chi\eta$ to $\sigma\tilde{\omega}\mu\alpha$). Human life, in its habitual form, is confined to the space left open between these two standpoints. It is a mixture of 'psychic' and non-'psychic' (or somatic) elements, whose inner composition is open to variation.

Soul and body do not stand, therefore, for different cognitive modes, but for different degrees of the same 'psychic' element. Whereas the soul represents a plenum of cognitive awareness, or a zero of cognitive hindrance, the body represents a zero of cognitive awareness, or a plenum of cognitive hindrance. But despite the soul's formal protagonism, its true or authentic meaning remains out of sight. Although Plato's description of the body-soul compound is often grounded in a description of the soul in its pure form, counterposed to the impurity of its fallen, embodied form, this kind of account is necessarily allegorical. Since none of us is or was ever a 'pure soul', our understanding of the scale's highest term is bound to be thoroughly speculative. Not only do we lack the kind of insight ascribed to a pure soul, we are also incapable of anticipating what it would be like to have it. Whenever we think of a purely 'psychic' standpoint, we do so negatively, by mentally detaching ourselves from our body and trying to imagine what would be left once our bodily needs and constraints were suspended.

This kind of limitation also applies, for similar reasons, to the other end of the scale. Since none of us is or was ever only a body, our understanding of the scale's lowest term is also purely speculative. Our cognitive access to a life devoid of cognitive access entails a performative contradiction, which we usually bypass by means of a mental fiction. Whenever we think of a 'pure body', we do so negatively, by mentally discarding our powers of cognition and trying to imagine what we would be left with. This is the kind of procedure we resort to when picturing our own death, which we do by stepping outside of ourselves, as it were, and placing our dead self in front of us. But this exercise is only possible, of course, because we remain conscious of ourselves throughout the whole process, no less alive than we were before.

³¹ *Phd* 67d1–2

³² *Phd* 82e1–4

³³ *Phd* 83d4–6

In short, all our cognitive representations (and *a fortiori* the representation of a purely psychic and a purely somatic standpoint) are both psychic *and* somatic, viz. psychic *and* non-psychic. In the body-soul compound, the compound itself has priority over what is compounded, so that the identification of its ingredients is always an immanent procedure. But Plato's aim is not simply to bring out the constitutive nature of this synthesis. He also points out that this compound is not peaceful or harmonious, but a source of tension and unease. In its usual, hindered form, the soul is the target of two different forces, pulling in opposite directions. On the one hand, it is naturally attached to its pure form. On the other, it is 'nailed' to the body, burdened with its needs and desires. Given this contradiction, the soul is not at home in the body. Imprisoned and denatured, expelled from its original element, it still holds on to the *memory* of it. In the physical world, the soul's primitive grandeur lives on in the form a faint but inextinguishable echo—an echo that corresponds, precisely, to the notion of truth.

We have thus arrived at the central element of Plato's description. The soul, deprived of its immediate communion with truth, is bound to look for it in the bodily realm. And herein lies the primary source of our usual concern for the truthfulness of our beliefs. Insofar as every human idea is born within the soul, it is accompanied by a natural attachment to truth. But whereas for a pure soul reality and true reality are the same thing, for a composite standpoint such as ours the notion of truth is a formal notion, lacking an immediate or self-evident content. There is, in our usual standpoint, an *excess* or a *disproportion* with regard to truth that turns every idea or judgement into a possible *version* of truth, or an *applicant* to the status of truth. Since truth is not automatic or transparent, it is bound to be translated into specific theses about reality, which may or may not be true.

In addition, this attachment to truth is not merely gnoseological, but practical and self-interested. Whereas a pure soul knows everything, including what is best or most advantageous, for a composite standpoint like ours the greatest good is not immediately apparent. It is also the object of a thesis, or of several different theses, whose content is bound to change with time and circumstances. The expectation of truthfulness underlying all our judgements and decisions is also accompanied by a belief in their ability to advance our interests. And just as the former may turn out to be groundless, so too is the latter open to refutation.

This general description brings us back to Plato's conception of *δοξάζειν*. Since the soul is always attached to the notion of truth, every act of cognition sets forth a specific version of things, usually assumed to be true. Every judgement is already a doxastic judgement, and this general rule is valid even in the lowest stages of cognitive awareness. Although none of us is ever a purely intellectual being, our bodily needs and desires are never so strong as to obliterate our natural concern with truth. Even when our soul falls prey to the most violent passions, or the most pressing physical demands, our enjoyment and pain are still dependent on the validation of a specific version of things. In these situations, the soul is still called

upon to assent to 'the things . . . which the body says are true.'³⁴ And though this cognitive approval is often automatic and unconscious, it is still a necessary condition for pain and enjoyment to be perceived as such.

Out of the various passages one might choose to illustrate the Platonic contrast between *σῶμα* and *ψυχή*, the *Phaedrus*' depiction of the soul as a flying chariot is arguably the most suggestive.³⁵ In this well-known passage, Plato sums up most of the issues mentioned above in a poetic, but nonetheless remarkably accurate manner. Although this is not the place for a thorough examination of this text, let us recall the main elements of Socrates' description. He starts by likening the soul's inner structure to a chariot guided by a charioteer and pulled by a pair of winged horses. This chariot is similar to the one that carries the gods across the heavens, but with a few important differences: while the charioteer and the horses of the gods 'are all good and of good descent',³⁶ those of the human soul are mixed. One of the horses, 'noble and of noble breed',³⁷ pulls the chariot upwards, towards the dwelling-place of the gods; the other horse, 'quite the opposite in breed and character',³⁸ persists in dragging it downwards, to the bustle of concrete pleasures and mundane affairs. Therefore, while the journey of the gods is safe and easy, the journey of human life is dangerous and difficult, prone to accidents and changes of course.

This description suggests, from the very beginning, three important ideas. First, the human soul, in its usual form, is a decayed or downgraded soul. Although it is naturally akin to the world inhabited by the gods, it fails to attain its degree of perfection. It is modelled after an ideal it is unable to fulfil. Second, the soul is not simple, but complex and self-divided. It is the battling ground of two opposite forces, represented by the horses' different breedings. Third, due to this self-division, the soul's nature is not fixed, but variable. Depending on the skill of the charioteer, and on the strength and speed of each horse, its journey can be very different and lead to very different places. These three ideas are all part of the definition of *ψυχή* summed up in the previous pages. First, the soul's usual form is indeed a long way from its essential or potential perfection. Instead of a true and transparent glance on the whole of reality, it is normally confined to a restricted cognitive perspective, grounded in a complex system of beliefs. Second, the soul is indeed self-divided, torn between an original urge for truth and clairvoyance and a practical urge for self-preservation, self-orientation and enjoyment. Third, depending on how strongly each of these tendencies is felt, the soul is indeed open to change. Its nature can vary according to a cognitive scale ranging from a very small degree to a very large degree of somatic hindrance.

³⁴ *Phd* 83d6

³⁵ See *Phdr* 246a–257a.

³⁶ *Phdr* 246a7–8

³⁷ *Phdr* 246b2–3

³⁸ *Phdr* 246b3

At first sight, the opposition between the noble and the rebel horse might be interpreted as a simple contrast between the rational element in human life, which leads one to favour wisdom and prudence over pleasure and gratification, and a more immediate or rash attitude, leading to the opposite behaviour. But although this kind of reading is plausible, it fails to capture the various subtleties of the allegory. Recall, firstly, that the whole description is about the soul and not about the body. Each allegorical element represents, therefore, a specific part or function of the soul. The charioteer, as the ruling element, symbolizes the soul's most primary function, which is that of granting access to reality. The charioteer represents not only cognition as such, but also the possibility of choosing different cognitive contents, as well as the need to settle for one in particular. The horses, on the other hand, symbolize the psychic and somatic tendencies within the soul, that is, the soul's choice between freeing itself from the interference of the body or 'nailing' itself further to it.

Regardless of which is dominant, both alternatives entail the endorsement of specific versions of reality, whose truth-value is determined by the entire soul. Once more, Plato is intent on showing that every act of cognition, however immediate or instinctive, presupposes a specific set of beliefs assumed to be true. And this is made clear by the fact that the unruly horse, although unwilling to cooperate with the charioteer and the noble horse, never acts on its own. It is bound to the chariot and forced to drag it along towards the things it prefers, so that the entire soul is led to regard those things as truthful and desirable. If the unruly horse proves stronger or more obstinate than the noble one, and if its jerks become too violent and impossible to restrain, the charioteer and the noble horse 'go forward with him, yielding and agreeing to do his bidding'.³⁹ In other words, even amid the most primitive bestiality or the most unbridled physical lust, when bodily needs gain the upper hand and our thoughts and actions seem oblivious to the pursuit of truth—even then it is the soul that renders pleasure pleasurable, pain painful, enjoyment enjoyable, and so on.

The soul, not the body, is the ultimate source of the body's impressions, feelings and affections. The body only 'seems to be self-moving, because of the power of the soul within it.'⁴⁰ This conclusion plays a major role in Plato's analyses of pleasure and desire, developed at length in dialogues like the *Philebus* or the *Gorgias*. In these texts, he aims to discredit the idea that pleasures are good in themselves or that pains are evil in themselves. And he does so precisely by arguing that pleasure and pain are never self-evident, but always dependent on the specific course taken by the soul's chariot. Anticipating one of Stoicism's main themes, Plato argues that our usual judgements concerning pleasure and pain are more flexible than we tend to suppose, allowing for a significant degree of variation.

This brings us to a second and crucial element of the allegory, namely the contrast between the ascending and the descending movement of the soul's chariot. Instead of picturing a vehicle travelling on wheels, along a road or a track, Plato describes an Olympian carriage, pulled by flying horses. This peculiarity conditions the whole

³⁹ *Phdr* 254b2–3

⁴⁰ *Phdr* 246c4

allegorical setting, subordinating all its elements to the alternative between moving forward and/or upwards, propelled by the horse's wings, or falling down, due to the force of gravity.⁴¹

To begin with, this alternative reiterates the idea of a soul confined to a downgraded version of itself, burdened by a calling it cannot fulfil. Not only is the soul's chariot modelled on a divine chariot, meant for perfect or omniscient beings, but its horses are also special horses, similar to those of the gods. Their wings suggest that their natural habitat is not the earth down below, but the sky up above—that they are destined to travel upwards rather than downwards. And this suggestion is made clearer in Plato's subsequent description of the soul's wings and of how they can either shrink into mere stumps or develop into full-fledged organs.⁴² At stake in this poetical (and comical) passage is once more a cognitive scale, measuring the degree of hindrance endured by the soul in its embodied form. In Plato's words, 'the natural function of the wing is to soar upwards and carry that which is heavy up to the place where dwells the race of the gods.'⁴³ Therefore, when a soul 'is perfect and fully winged, it mounts upward and governs the whole world; but the soul which has lost its wings is borne along until it gets hold of something solid, when it settles down, taking upon itself an earthly body.'⁴⁴ Yet although the embodied soul is fated to 'lose its wings' (πτερορρηῖν), it is never *fully* wingless (ἄπτερος): its wings are always present, even if only as mere stumps, because the soul is permanently affected by the memory of its divine provenance; or, again, because whatever its position within the cognitive scale, it already implies some kind of relationship, however remote, with the scale's highest term.

This point is crucial: however weighty or earthbound, the soul is always haunted by a perfect, or perfectly winged version of itself. Indeed, this haunting does not refer simply to a *better* version of itself, but for the *absolute best*, the *best there is*, the *best thing conceivable*. And this superlative does not indicate merely a plenum of cognitive insight, but also a plenum of satisfaction and self-fulfilment. Since the soul, in its purest or highest form, knows everything there is to know, it also knows what is best, and leads the best possible life. Its original dwelling place is nothing less than ὁ ὑπερουράνιος τόπος, i.e. the very top of the celestial vault, where the gods are said to dwell. But this image may once again be misleading: when we picture this divine realm, we tend to place it somewhere above the clouds, or above the sky, or above the universe—at a great, but finite distance from the human realm. However, this ontological summit, precisely because it stands for the *absolute best*, can never be placed at a finite distance. Although it is always somehow 'close at

⁴¹ The contrast between an upward and a downward movement is also a structuring theme in the allegory of the cave and in the *Phaedo's* final cosmological myth. Along with the *horizontal* contrast between physical closeness and physical distance, this *vertical* contrast is one of Plato's preferred devices to convey human life's ambiguous condition. See Sect. 4.2, fn. 31.

⁴² *Phd* 246b6ff.

⁴³ *Phdr* 246d6–7

⁴⁴ *Phdr* 246b7–c4

hand', in the form of a longing or a haunting, it is by definition *thoroughly inaccessible and out of reach*. If the *ὑπερουράνιος τόπος* were to correspond to a definite place, with definite qualities and a definite appearance, it would cease to be the best thing conceivable. The soul would be able to conceive something *even better*, even if only in a negative way. Accordingly, the superlative determination which the soul aims at is in fact the *superlativized form of a comparative determination*, continually open to higher forms of meaning. It is not simply *τὸ βέλτιστον*, or *τὸ κάλλιστον*, or *τὸ ἀληθέστατον*, but a superlative form of *βέλτιον*, a superlative form of *κάλλιον*, or a superlative form of *ἀληθέστερον*.

Finally, let us turn our gaze downwards. Apart from the soul's wings, and their varying degrees of development, the chariot's movement is also conditioned by the force of gravity. And this simple fact is also of great importance. In order to keep the chariot moving, the horses must be able to counter its natural tendency to fall. But while the horse of noble breed, determined to pull the chariot upwards, is forced to exert an active effort, the unruly horse is helped downwards by the force of its own weight. The automatic nature of gravity, as opposed to the artificial nature of the soul's ascending flight, illustrates the radical difference in appeal of the psychic and the somatic components of human life. Whereas the latter relates to concrete facts and decisions, the former amounts to a much less evident possibility, usually perceived as vague and metaphysical, more suited to philosophical enquiries than to the practical concerns of everyday life. Bodily truth is the truth we (literally) fall back on, the one we most readily recognize and accept. It is the cognitive standard we adhere to by default. 'Psychic' truth, by contrast, requires a voluntary shift of perspective—a 'soaring upwards', away from the body and its usual concerns.

Plato's allegory reveals a very peculiar crossroads, whose alternative routes are not truly equivalent, nor equally visible, nor equally comfortable or easy to tread. Yet both of them condition, in varying degrees, the kind of life we lead, the sort of things we value and the aims and projects we favour and pursue. They represent different attitudes towards life, knowledge, and happiness, whose different existential translations will be considered in more detail in the next chapter.

Chapter 3

Masters, Slaves and Philosophers



Abstract Building on the contrast between a self-interested (or ‘somatic’) and a selfless (or ‘psychic’) attitude towards truth and science, Plato distinguishes between two different kinds of life: the philosophical life pursued by Socrates and the doxastic life led by most human beings. This chapter examines these two alternatives and their connection to Plato’s conception of freedom.

In Sect. 3.1, starting from the contrast between rhetors and philosophers drawn in the *Theaetetus* (172c–173b), I discuss Socrates’ claim that rhetors, when compared to philosophers, are like ‘slaves in breeding compared with freemen.’ In Sect. 3.2, I analyse the counter-claim, made by Callicles in the *Gorgias*, that the philosophical life pursued by Socrates is ‘more befitting a slave’. In direct opposition to Socrates, Callicles embodies an extreme and self-conscious form of *doxazein*, which prefigures Nietzsche’s repudiation of Plato’s philosophical ideal. Finally, in Sect. 3.2, I delve more deeply into the debate between Socrates and Callicles, which offers one of Plato’s most explicit discussions of the cognitive foundations of human freedom.

Keywords Plato · Socrates · Freedom · *Gorgias* · Callicles · Masters · Slaves

Throughout Plato’s dialogues, Socrates often distinguishes between two different ways of life, based on two different kinds of knowledge. In the first case, human beings are focused on the practical efficiency of their views and judgements. Their interest in truth is superficial and the standard whereby they accept or reject a given idea or belief is not very demanding. In the second case, this standard is raised and truth is valued more highly. The priority of efficiency over accuracy is reversed and one’s beliefs are subjected to a closer examination.

In the previous chapter, the first of these attitudes was described with the aid of the *Philebus*. I have shown, following Plato, that the observer imagined by Socrates is not a mere spectator, but a self-interested agent, committed to a complex system of assumptions, aims and aspirations. When faced with a distant object, his main concern is not to get to the bottom of what he sees, but simply to gather as much information as necessary in order to settle for a determinate version of things. Since

the identity of the object is determined by his practical situation, its meaning is always provisional. As the observer's situation changes, so too does the object's definition.

This example illustrates what goes on with most of us throughout most of our lives. Our account of reality is also grounded in a myriad of practical concerns, whose meaning and importance are continually revised. But although we do not value truth as such, our cognitive standpoint is dependent on the notion of truth. And indeed just as much so as a purely epistemic standpoint. Though we do not often think about truth and error, we do wish our actions to be successful and our goals to be reached. And for that to happen our judgements about reality must be reliable and reveal to us the world as it truly is.

Plato's dialectical charges are usually directed at this very demand. By pointing out the unreliability of our ideas and judgements, he means to highlight our usual incapacity to secure our own interests. If the efficiency we strive for is dependent on the accuracy of our beliefs, and if the latter is open to criticism, our life choices are likewise open to criticism. For Plato, the solution to this problem requires the adoption a new mode of cognition, guided by a different relationship with truth.

3.1 Rhetors and Philosophers

Halfway through the *Theaetetus*, as Socrates and Theodorus discuss the exact definitions of δόξα and ἐπιστήμη, their debate is momentarily sidetracked by a secondary, albeit closely related discussion. From 172c on, Socrates compares two different kinds of life, based on two different attitudes towards knowledge and truth. The first is ascribed to lawyers and public speakers and the second to philosophers, or 'lovers of knowledge'.

Let us focus on the first these two alternatives. Though the passage in question is somewhat long, it is important to consider Plato's exact wording:

SOCRATES: Those who have knocked about in courts and the like from their youth up seem to me, when compared with those who have been brought up in philosophy and similar pursuits, to be as slaves in breeding compared with freemen.

THEODORUS: In what way is this the case?

SOCRATES: In this way: the latter always have . . . leisure, and they talk at their leisure in peace; just as we are now taking up argument after argument, already beginning a third, so can they, if as in our case, the new one pleases them better than that in which they are engaged; and they do not care at all whether their talk is long or short, if only they attain the truth. But the men of the other sort are always in a hurry—for the water flowing through the water-clock urges them on—and the other party in the suit does not permit them to talk about anything they please, but stands over them exercising the law's compulsion by reading the brief, from which no deviation is allowed (this is called the affidavit); and their discourse is always about a fellow slave and is addressed to a master who sits there holding some case or other in his hands; and the contests never run an indefinite course, but are always directed to the point at issue, and often the race is for the defendant's life.¹

¹Th 172c8–e7

At first glance, Socrates' description is rather specific: his mention of courts of law (δικαστήρια), public contests (ἀγῶνες), an opposing party (ἀντίδικος) and an affidavit (ἀντωμοσία) seems to leave no doubt as to the legal nature of his account. The rhetors or public speakers he refers to are lawyers or litigators, engaged in a judicial hearing, and what is being highlighted is the difficulty of arriving at true judgements during a legal dispute, inasmuch as public litigations do not allow lawyers the time or detachment needed to make substantiated arguments. Since they are always in a hurry, and since their decisions can result in the life or death of their defendants, they can hardly remain impartial and serene.

This permanent tension is bound to take its toll on the litigators' character and to shape it in negative ways:

As a result of all this, the speakers become tense and shrewd; they know how to wheedle their master with words and gain his favour by acts; but in their souls they become small and warped. For they have been deprived of growth and straightforwardness and independence by the slavery they have endured from their youth up, for this forces them to do crooked acts by putting a great burden of fears and dangers upon their souls while these are still tender; and since they cannot bear this burden with uprightness and truth, they turn forthwith to deceit and to requiting wrong with wrong, so that they become greatly bent and stunted. Consequently they pass from youth to manhood with no soundness of mind in them, but they think they have become clever and wise.²

In addition, Socrates points out that the aim of legal arguments is not so much to be truthful as to be efficient, and that the job of litigators is not so much an art of truth as an art of persuasion. In order to dominate their opponents and win over their audience, litigators rely on artifice, demagoguery and increasingly sophisticated forms of deceit. Therefore, as they become better at their job, 'their souls become small and warped' and 'they pass from youth to manhood with no soundness of mind.'

By contrasting the life of public speakers and litigators with the life of 'those who have been brought up in philosophy', Socrates introduces a familiar Platonic theme. Despite his focus on legal justice and the specific proceedings of an Athenian court of law, his description turns up in connection with a debate concerning the teachings of Protagoras of Abdera, the most famous and the most conceited of all Greek sophists. Just like Protagoras—as well as Gorgias, Prodicus, Hippias or Thrasymachus, among others—the rhetors mentioned in the *Theaetetus* consider themselves 'clever and wise' and 'know how to wheedle their master with words and gain his favour by acts.' Their distinguishing feature is not their wisdom or virtue, but their ability to persuade others that they are wise and virtuous.

As in so many of the dialogues, Plato is alluding to the fundamental distinction between philosophy's primary concern with wisdom, or σοφία, and rhetoric's primary concern with persuasion and public approval. He is pointing out that the eloquence exhibited by most professional speakers is not indicative of the truthfulness or cogency of their arguments, but merely of the ability to pass them off as true

²Th 173a1–b3

and cogent. Whereas philosophers are concerned with the clarity and consistency of each idea or judgement, rhetors and sophists are merely focused on the ‘rough and tumble of arguments’³. Like the litigators mentioned above, ‘they do not care what the truth is in the matters they are discussing, but are eager only to make their own views seem true to their hearers.’⁴

At first glance, then, Socrates’ digression in the *Theaetetus* can be read as yet another charge against the widespread popularity of rhetoric and sophistry.⁵ And this is all the more plausible since the role of public speakers in Athenian society during Plato’s lifetime was not confined to philosophical or epistemological debates. Their political and social influence was considerable and their services were requested by many citizens at numerous junctures of personal and public life. Rhetors and sophists acted as intermediaries in job applications, matrimonies and lawsuits⁶, and it is the latter capacity that Plato highlights in the passage we are considering.

However, this reading only reveals a small part of what goes on in the text. Despite this classical topos, Socrates’ distinction between ‘those who have knocked about in courts and the like’ and ‘those who have been brought up in philosophy’ must not be interpreted too strictly. This contrast also speaks to a wider and more profound cleavage, separating two radically different modes of being. Socrates’ condemnation of the hasty and self-interested attitude that inspires most legal decisions is not aimed solely at litigators or rhetors, nor at a specific kind of elocution. He is also suggesting that all human beings, as regards their usual cognitive perspective, are rhetors or litigators in their own right.

To appreciate the true reach of Plato’s description, let us examine its main elements. He starts by mentioning the problem of time and the fact that litigators cannot reflect at leisure, but are forced to make sudden and often unsubstantiated decisions. Whilst philosophers can consider several arguments, study each one in detail and avoid reaching a conclusion until they are absolutely satisfied, the rhetors mentioned by Socrates are not as free. The strict rules that govern legal proceedings prevent them from being thorough in their assessments and the pressing nature of legal disputes calls for speedy resolutions. In court, the clock is always ticking.

³So 268a2

⁴Phd 91a3–6

⁵The words ῥήτωρ and σοφιστής may be used as synonyms, but they differ in important respects. The former, used in the *Theaetetus*, applies to a public speaker in all his different capacities, while the latter is derived from the adjective σοφός (‘wise’, ‘skilled’) and the verb σοφίζομαι (‘to make wise’, ‘to instruct’), and refers to someone who excels in a determinate activity. From the 5th century B.C., however, the word *sophist* acquired a more specific meaning, related to the rise of the sophistical movement. The sophists of Plato’s day were itinerant teachers whose discourses were highly praised and whose lessons were often very expensive. Simultaneously, however, the word *sophist* also acquired a negative tinge, explored in Plato’s writings and still present in the modern terms *sophism* and *sophistry*. See Gomperz, *Griechische Denker*, vol. 1, 345ff., and Guthrie, *A History of Greek Philosophy*, 27–34.

⁶See Philonenko, *Leçons Platoniciennes*, 81; or Marrou, *Histoire de l’éducation dans l’antiquité*, 95.

The timing of a decision may be as important as its content, and a good call made too late can be worse than a bad one made on time.

Furthermore, Plato highlights the self-interested nature of legal decisions. Whilst philosophers take a step back to observe and judge, looking at each issue from different angles, rhetors and litigators are not as circumspect. Given their interest in their own success and in that of their clients, as well as their usual longing for the crowd's approval, their decisions are never selfless or neutral, but always 'tense' and 'shrewd'. For rhetors and litigators, it is not indifferent to decide or not to decide, to decide this way or that way, to decide sooner rather than later, etc. Unlike mere observers, they are continually forced to take a stand, and their decisions lead to real and inescapable consequences.

This diagnosis is now becoming more familiar. Recalling our previous discussion, the decisions now linked to the activity of litigators are very similar to those Plato calls δόξαι. And their self-interested nature, grounded in a permanent urge for practical success and orientation, is also evocative of Plato's characterization of δοξάζειν. His criticism of the deliberative method employed by legal orators can be read, therefore, as a criticism of human cognition in general. The rhetors mentioned in the text stand for all of us, throughout most of our lives: just like litigators, we are wrapped up in our own vital cause; the clock urges us on, as it does all litigators, forcing us to choose between different judgements, deeds and life paths; just like litigators, we are used to sacrificing truth and accuracy to practical efficiency; and hence, just like litigators, we are continually led to rely on rash and unsubstantiated decisions.

All this would not be as serious were we aware of it. But the point is precisely that we are not. To be sure, there is a conscious dimension to the deceitful nature of rhetoric. Rhetors and litigators are cunning and shrewd, well aware of the ruses they resort to in their speeches. Plato's attack on rhetoric is thus an attack on their usual dishonesty, on their 'crooked acts' and their tendency to 'requite wrong with wrong'. But his criticism also refers to a more profound cognitive tendency, of which the rhetors' conscious dishonesty is only the most visible example. He accuses litigators (and ourselves) of having 'no soundness of mind' due to their (and our) tendency to overlook the inherent limitations of their (our) decisions and to mistake the practical convenience of a given judgement for its actual truth. In this cognitive regime, each new decision reinforces the litigator's (and our) illusion of having 'become clever and wise', when in fact it has only helped establish a provisional and ultimately unwarranted version of things.⁷

⁷The term 'rhetorical' might be used to stress the overly theoretical nature of a given discourse—as when one says that a question is 'merely rhetorical'. But Plato's critique of rhetoric is meant to show the opposite. Rhetors are above all practical men, concerned with practical truths, and although they are capable of discussing highly abstract epistemological issues, their ultimate goal is to contradict their opponents and win over their audience. It is based on this practical dimension that our own lives can be deemed 'rhetorical': like Plato's ῥήτορες, we are primarily focused on practical rather than theoretical concerns and willing to sacrifice the epistemic accuracy of our beliefs to their practical efficiency.

We find here yet another instance of a typically Platonic procedure: in spite of focusing on the activity of orators and litigators, Plato is making the much more general point that all human beings usually think they know what they do not and base their entire lives on ‘facts’ which are in truth mere beliefs. This general misconception affects all human ideas and judgements, and it also extends to the tricks and ruses favoured by rhetoric. Since rhetors are also prey to this peculiar form of ignorance, the tricks with which they deceive others are ultimately also a form of self-deception. Whenever they lead others into thinking something is true, or good, or advantageous, they do so based on their own judgement of what is true, or good, or advantageous. But since the latter is just as doxastic as every other judgement, its truth-value is just as open to refutation.

After this brief description, Socrates goes on to characterize the life of philosophers. The first difference between rhetors and philosophers, already mentioned, is their relationship with time: whereas the former are ‘always in a hurry’, with the water-clock ‘urging them on’, philosophers ‘talk at their leisure in peace’ and ‘do not care at all whether their talk is long or short, if only they attain the truth.’ Second, whereas rhetors are professional orators, masters of eloquence and elocution, philosophers ‘appear ridiculous when they enter the courts of law as speakers’.⁸ Their quest for truth leads them away from the concrete concerns of everyday life and into the world of thought, making them ignorant of all things practical and political. To illustrate this peculiar kind of alienation, Socrates evokes Thales of Miletus, who is said to have fallen into a pit while looking upwards to study the stars. Tradition has it that a Thracian servant girl, having watched the incident, made fun of the philosopher on account of his being ‘so eager to know the things in the sky that he could not see what was there before him at his very feet.’⁹

Through this well-known anecdote, Thales emerges as the archetypal philosopher, clumsy and absent-minded, absorbed by theoretical problems and oblivious to concrete daily concerns. According to Socrates, what sets Thales and most philosophers apart from ordinary human beings is not a lack of interest in human life—quite the contrary: precisely because they are unusually drawn to human nature and intrigued by what it means for human beings to think and act the way they do, they tend to disregard life’s contingencies and focus on its most general or essential aspects. Although Thales is ‘ignorant of what he is doing’ and ‘hardly knows whether he is a human being’, he exerts himself to determine ‘what a human being is and what is proper for such a nature to bear different from any other.’¹⁰

The list of Thales’ descendants is long and varied. His closest successor is Archimedes, whose excitement over the discovery of a fundamental law of physics is said to have made him jump out of his bath and take to the streets of Syracuse

⁸Th 172c5–6. In the *Apology* (17a–18a), Socrates starts his legal self-defence by contrasting his manner of speech, spontaneous and plain, with the ‘speeches finely tricked out with words and phrases’ that are common in public gatherings and courts of law.

⁹Th 174a6–8

¹⁰Th 174b

naked, eager to communicate his findings.¹¹ His lack of common sense is also brought out in a well-known anecdote concerning his death, recounted by Plutarch: during the Roman siege of Syracuse, Archimedes was absorbed in the contemplation of a mathematical problem and did not notice the military manoeuvres unfolding around him. When a soldier came up to him, with orders to take him to the general in charge of the Roman troops, Archimedes refused to break off his reflections until he had found a solution to his problem, and ended up being killed.¹²

But whereas Archimedes' death is tragic, the usual remoteness of philosophers has also been explored to comic effect. One of the best examples can be found in Aristophanes' *Clouds*, where Socrates and his pupils are depicted as the fanatic enthusiasts of a ludicrous philosophical sect, whose main practices consist in walking on air, contemplating the sun and finding out the truth about all sorts of 'celestial matters'.¹³ The satire caricatures both the absent-mindedness of philosophers and their utter disinterest in (and disdain for) concrete everyday life. The objects of their studies range from minute biological details, such as the physiology of small insects, to fantastic cosmological theories, concerning the origin and structure of the universe. In all of these cases, however, their conclusions are of little or no practical value. Their alleged wisdom is humorously contrasted with their lack of common sense.¹⁴

The definition of philosophy as a form of alienation is a recurrent theme in Plato's writings, where Socrates is also portrayed as an outsider whose ideas, concerns and priorities are very different from those of regular human beings. On the one hand, his enquiries into the structure and scope of human cognition call into question what most people take for granted. His insistence on reason's intrinsic contradictions leads him to challenge not only the social, religious and political tenets of Athenian society, but also the very foundations in which human existence is usually grounded. As the dialogues unfold, philosophy is increasingly revealed as an *absolutist* enterprise, whose universal vocation is incompatible with ordinary reason. On the other hand, however, Plato seeks to show that the alienation produced by philosophy is only perceived as such from a non-philosophical point of view. Socrates appears eccentric and unreasonable precisely because his ideas are judged from a standpoint he himself deems unreasonable. If human reason were to assume a different form and favour different ideas and concerns, this judgement might be radically altered.

¹¹ This anecdote first appeared in Vitruvius' *De Architectura*.

¹² See Plutarch, *Plutarch's Lives*, chapter 19, §§4–6. In Plato's dialogues, Socrates' keen interest in theoretical matters also leads him to forget himself, other people and his immediate surroundings. See Sy 174d4–175b3 or 220c1–d5.

¹³ Aristophanes, *Clouds*, l. 37. This portrayal of Socrates is explicitly mentioned in *Ap* 19c.

¹⁴ See Aristophanes, fr. 672 K. I 557; *Birds*, ll. 1553ff. and 1282ff.; *Frogs*, ll. 1491ff. In the *Clouds*, however, Aristophanes does not focus solely on the philosopher's absent-mindedness. His tone, as Nietzsche puts it, is 'half outraged, half scornful' (BT, 65), and he does not hesitate to accuse Socrates of being 'all affectation and fuss' (l. 299). Interestingly, Aristophanes' charges echo Socrates' own charges against the rhetors and sophists of his day. Socratic philosophy is portrayed not only as devoid of practical value, but also as downright fraudulent. The Aristophanic Socrates is not naïve, but a charlatan who seeks to profit from his supposed wisdom.

3.2 Socrates Versus Callicles

In the second half of the *Gorgias*, Socrates engages in a heated debate with Callicles about the merits of rhetoric and philosophy. During this long exchange, Callicles underlines the eccentricity of philosophers in much the same way as Socrates does in the *Theaetetus*. But whereas Socrates' account culminates in a redeeming moment, where philosophy is presented as a real alternative—and indeed as the only true alternative—to the life of rhetors and litigators, Callicles' description does not. He believes that the practice of philosophy must only be encouraged at an early age, and even then only to a limited extent. In adult life, philosophy is not only useless, but also harmful, and even dangerous:

However well endowed one may be, if one philosophizes far on into life, one must needs find oneself ignorant of everything that ought to be familiar to the man who would be a thorough gentleman and make a good figure in the world. For such people are shown to be ignorant of the laws of their city, and of the terms which have to be used in negotiating agreements with their fellows in private or in public affairs, and of human pleasures and desires; and, in short, to be utterly inexperienced in men's characters. So when they enter upon any private or public business they make themselves ridiculous.¹⁵

In his long and scornful speech, Callicles concedes that 'it is a fine thing to partake of philosophy just for the sake of education', and hence that 'it is no disgrace for a lad to follow it.'¹⁶ Philosophy is a suitable occupation for young men, and the mark of a free or liberal mind (ἐλευθερος).¹⁷ But when schooling gives way to real life, and one's decisions become serious and binding, philosophy is no longer beneficial. In its most innocent version, it is a source of ridicule for those who pursue it. Inexperienced in practical matters and public affairs, they are bound to be made fun of—in the same way that Thales is made fun of by the Thracian girl, or Socrates by Aristophanes. In its most dangerous version, however, philosophy may become a serious liability. Since philosophers are usually naïve creatures, unable or unwilling to take an active part in social and political life, they are also less able to defend their interests and to protect themselves against other people. Consequently, whenever their words are judged incommensurable to others, they are liable 'to be stripped by [their] enemies of all [their] substance, and to live in [their] city as absolute outcast [s].'¹⁸

With these two worries in mind, Callicles alternates between a patronizing attitude and a more ominous, quasi-menacing tone. At first, he describes Socrates' life-long engagement with philosophy as ridiculous and unmanly, 'more befitting a slave' than a gentleman.¹⁹ He likens his obstinacy to a 'boyish travesty', unworthy of his noble nature and rendered grotesque by his advanced age. Afterwards,

¹⁵ *Go* 484c7–d7

¹⁶ *Go* 485a

¹⁷ *Go* 485c

¹⁸ *Go* 486b7–c2

¹⁹ *Go* 485a

considering Socrates' usual insolence and the animosity generated by his enquiries, Callicles' mockery gives way to a more serious admonition, complete with proleptic allusions to Socrates' forthcoming trial and execution.²⁰

Callicles' criticism of philosophy and philosophers stands out for its vehemence, but it is by no means an exception within the Platonic corpus. Depending on the interlocutor and the context, Socrates' views are often met with different kinds of resistance: they are the target of perplexity, whenever they seem too absurd or contrary to common sense; ill will, whenever they hurt his opponent's pride or self-esteem; indifference, when they are deemed irrelevant for the successful running of everyday life; or reproach, when they are perceived as intrusive or dangerous.²¹ Again, what is interesting about this catalogue is not simply what these reactions say about Socrates' views, but what they tell us about ourselves. When reading Plato's dialogues, we usually make two automatic assumptions, which condition our entire understanding of the text: first, we tend to assume that the views held by Socrates and his interlocutors are equally clear or intelligible; second, we tend to liken our own standpoint as readers to that of a referee, whose role is to weigh the arguments on both sides and decide on their merits.²² These assumptions, however, conceal the true complexity of Plato's strategy.

To begin with, the dialogues are never merely two-way conversations, with Socrates on one side and his interlocutors on the other, but *three-way* conversations, where the reader also plays an important part. Plato is continually bringing us into the game, so to speak, and playing on our assumptions and expectations. His arguments are often built in such a way as to address a specific prejudice of ours, to elicit a specific reaction on our part or to catch us off guard on a specific subject. Furthermore, the dialogues are never simple contests between symmetrical or equivalent standpoints. Although we tend to put ourselves in the shoes of both contenders, their worldviews are not equally familiar or accessible to us. And this is so, once again, because we are never mere spectators or referees: when reading the dialogues, we are already committed to a specific outlook on reality, and our understanding of each argument is therefore just as biased and self-interested as our understanding of anything else.

²⁰The tragic tone of *Go* 486a-b is worthy of Sophocles or Euripides: '[CALLICLES, addressing SOCRATES:] . . . if somebody should seize hold of you . . . and drag you off to prison, asserting that you were guilty of a wrong you had never done, you know you would be at a loss what to do with yourself, and would be all dizzy and agape without a word to say; and when you came up in court, though your accuser might be ever so paltry a rascal, you would have to die if he chose to claim death as your penalty.'

²¹For good examples of each of these attitudes, see *Go* 481b6–c4, *Pr* 333e–334a; *Eu* 11b4–5 and *Go* 485e1–496d1.

²²This only applies, of course, to dialogues where there is a disagreement between Socrates and his interlocutor(s). In some dialogues, Socrates has no opposition and his interlocutor's role is simply to provide the cue for his reasonings.

From where we stand, Plato's dialogues amount to the confrontation between a standpoint that is familiar to us, embodied by most of Socrates' opponents, and one that is not, usually embodied by Socrates himself. Although we may empathize with his views, and even share some of them, they are nonetheless the expression of a foreign world, made of new and strange possibilities. Confined as we are to our usual doxastic perspective, we tend to interpret Socrates' suggestions in a doxastic way, either by downplaying their subversive power or by construing them in a negative way, as a bridge or a pathway to some cognitive beyond. Accordingly, the different ways in which Socrates' interlocutors react to his ideas reflect our own instinctive reactions to them. Firstly, we are often perplexed at Socrates' suggestions and tend to regard them as too extreme. We may accept that some of our beliefs about reality are uncertain and unintelligible, but we are hardly prepared to accept that *all of them* are open to revision. This idea seems too fantastic to be believed and too contrary to common sense. Secondly, Socrates' suggestion also represents a hard blow to our cognitive self-esteem, and tends to be resisted in direct proportion to its degree of severity. Again, our cognitive powers may not be perfect, but they are certainly good enough to provide us with a reasonable understanding of what goes on around us. Thirdly, even if the previous assumptions lose their force, and we come to take Socrates' suggestions more seriously, we tend to minimize their effects by downplaying their importance. In other words, Socrates may be right on a theoretical level, but his denouncements do not prevent us from carrying on with our lives, pursuing our usual aims or recognizing the things that really matter to us. Finally, even when this last assumption is questioned, and we catch a glimpse of the actual seriousness of Socrates' charges, we can still protect ourselves by looking the other way. There may be indeed a constitutive problem with how we perceive reality, but why choose to dwell on it, at the risk of sacrificing our usual self-confidence? Why not simply carry on as usual, however unphilosophical that may be?

These reactions, instinctively shared by Socrates' opponents and ourselves, speak to the tenacity of the bond that binds us to our usual cognitive situation. Like Callicles, the litigators mentioned in the *Theaetetus* and the vast majority of people, we are naturally convinced that our usual perspective, despite all its imperfections, can reveal what is actually important in life, whereas philosophy's abstract discourses cannot. Unlike Socrates, we are concerned with *concrete* things: with things that count, whose actual existence cannot be disputed and whose importance every reasonable person will allow.²³ 'Reasonable people' are of course modelled after ourselves: they too share our preference for concrete things and our instinctive distrust of philosophy's counter-intuitive reasonings.

Now what is at stake in Plato's writings is precisely the possibility of reversing this state of affairs. If the dialogues are read as the opposition between a familiar standpoint and an unfamiliar one, the reader remains fundamentally unchanged.

²³ As Hegel will point out, 'a reception of this kind is usually the first reaction on the part of knowing to something unfamiliar; it resists it in order to save its own freedom and its own insight, its own authority, from the alien authority.' (PS, 35 / HW 3, 55).

With a familiar perspective to hang on to, consciousness is allowed to take refuge in its usual web of assumptions and to remain within the doxastic realm. But Plato aims to prevent this by turning what is familiar into a problem. As our beliefs and expectations are challenged, we are driven out of our usual cognitive element and led into a new and strange one, where the previous rules no longer apply. Whenever this operation is successful, our identification with Socrates' opponents loses its force and our attitude towards his ideas is transformed. Instead of mere eccentricities, they become for the first time serious possibilities, as relevant as our usual ideas and concerns.

This, then, is Socrates' primary aim: to induce his listeners to shed their skin and contemplate the possibility of a new cognitive order. This peculiar metamorphosis is compared by Alcibiades, near the end of the *Symposium*, to the 'ravishment' produced by the music of Marsyas, the satyr-god, or by the song with which the Sirens tried to lure Odysseus. In a passage that evokes Callicles' indictment of philosophy, Alcibiades points out the unique power of Socrates' words. Whilst other orators, albeit very talented, are often unable to capture their listeners' attention, Socrates is always sure to astound and entrance. As an example of this, Alcibiades speaks of his own experience and the 'strange effects' he felt in connection with Socrates' discourses:

When I listened to Pericles and other skilled orators I thought them eloquent, but I never felt anything like this; my spirit was not left in a tumult and had not to complain of my being in the condition of a common slave: whereas the influence of our Marsyas here has often thrown me into such a state that I thought my life not worth living on these terms. . . . For he compels me to admit that, sorely deficient as I am, I neglect myself while I attend to the affairs of Athens. So I withhold my ears perforce as from the Sirens, and make off as fast as I can, for fear I should go on sitting beside him till old age was upon me.²⁴

In this telling passage, philosophy is once again deemed incompatible with the concrete demands of social and political life. Like Callicles, Alcibiades accuses Socrates of diverting attention from the duties a responsible citizen 'ought to mind'. When one is led into the maze of philosophy, what was important and respectable no longer appears so and one's position in the world is suddenly at risk. By Alcibiades' own admission, only by fleeing Socrates and his charms was he able to think straight, value what he used to value and cherish the 'favours of the many'.²⁵ Although he knows Socrates' teachings to be important, and feels ashamed to neglect them, he is instinctively drawn to 'take a runaway's leave'²⁶ and preserve what is left of his public life.

This passage also brings up a theme that was already present both in Callicles' speech and in the *Theaetetus*' legal description: namely, the link between the conflicting lives proffered by Socrates and his opponents and the contrast between freedom and slavery. While Callicles considers the pursuit of philosophy 'befitting a

²⁴Sy 215e4–216a8.

²⁵Sy 216b5

²⁶Ibid.

slave' and urges Socrates to break his dishonourable habit, the latter characterizes the rushed lives of litigators as a form of slavery, and points to philosophy as a form of release. Likewise, Alcibiades recounts how his own life was revealed by Socrates to be that of a 'common slave', and how he had to flee his presence to regain his sense of freedom.

This motif brings us back to the correlation between freedom, power and knowledge, which lies at the centre of Plato's philosophy. Callicles' position is that of most 'reasonable people', and the one we naturally tend to endorse: from our point of view, philosophy is enslaving because it compromises the attainment of the very things we regard as liberating and empowering. If we cannot trust our usual account of reality, or meet the concrete challenges of everyday life, or attain the 'means, repute and other good things' praised by Callicles,²⁷ our freedom will indeed be very limited. Our philosophical qualms, if given free rein, will end up preventing us from getting ahead in life and sharing in its pleasures and rewards. Too naïve or conscientious to raise our voices and fight for our interests, we will be left at the mercy of others, exposed to all sorts of abuse.

For Callicles, for Plato's litigators and for all of us, freedom is dependent on the ability to get what we want. And what we want is nothing less than *the best there is*—or, at any rate, the best we can get at each given moment. Yet both for Callicles and for us the identification of what is best only allows for a limited degree of variation. Our specific life-goals may differ, but we all wish to have 'means, repute and other good things in plenty'. And although the value or desirability of certain things may be judged differently by each of us, there are other, more general things that no 'reasonable person' would refuse. Indeed, life's main difficulties are not usually about deciding what is best or worth having, but about the actual possibility of obtaining it. Most of our dilemmas are practical rather than theoretical or philosophical: they relate to the efficacy of our choices and not to their actual meaning. Since we usually know what we want, or what we would like to have if given the opportunity, philosophy is of limited interest to us. At the end of the day, we would all trade its abstract truths for the concrete possibility of having 'means, repute and other good things in plenty.'

For Plato, however, this attitude is already the product of an assumption. And like all assumptions, it is subject to being disproven. If our entire belief system is exposed to the possibility of error, its highest-order determination can also be mistaken. If the self-interested drive that fuels our choices is based on an automatic definition of the greatest good, all our aims and priorities may be in need of revision. And therefore, though we usually believe we are choosing and acting in our best interest, we might be doing just the opposite: the 'good things' we set our sights on may not be so good after all, and the freedom they are supposed to provide may not be as liberating as we think.

²⁷Go 486d1

3.3 Masters and Slaves

In the *Gorgias*, Socrates' debate with Callicles about the merits of philosophy is based on a fundamental question, 'which has the highest conceivable claims to the serious interest even of a person who has but little intelligence—namely, what course of life is best.'²⁸ While Socrates maintains that the definition of the best kind of life is not immediate or self-evident, Callicles, Gorgias and Polus are all convinced that the best things in life are not that difficult to identify. To be sure, each human being's idea of the best life is personal and likely to change with time and circumstances: while one person may spend his or her life eating, another may regard hunting as the most vital of occupations; and while the former may develop, with time, a passion for hunting, the latter may succumb, in turn, to a life of gluttony. But although the variety of life choices available to human beings is indeed endless, and continually open to revision, Socrates' opponents are nonetheless convinced that there are more general definitions of the best life which everyone can agree on.

The first one is provided by Gorgias, who links the attainment of the best possible life with the art of rhetoric.²⁹ In his eyes, rhetoric is superior to all the other arts (τέχναι) because its scope is virtually unlimited. While experienced doctors, for example, possess a special insight into the inner workings of the human body, and are thus more able to avert disease and suffering, they are usually as ignorant as most people with regard to the art of hunting. And while hunters are more skilled than other people on the hunting ground, and hence less vulnerable to malnutrition and death, they are usually as ignorant as most people with regard to medicine. And the same holds for all the other τέχναι: each grants a privileged insight into a specific aspect of reality, leaving the others more or less untouched.

For Gorgias, skilled rhetors differ from skilled doctors or hunters in that they are not limited to a specific kind of expertise. Since their eloquence enables them to talk doctors, hunters and other τεχνῖται into doing what they want, they can reap the benefits granted by other τέχναι without having to learn them. The actual knowledge about medicine and hunting is still a prerogative of doctors and hunters, but its empowering effect, which derives from the actual ability to interpret and change the course of events, is appropriated by rhetors and used in their favour. Moreover, skilled rhetors are more persuasive than doctors, hunters and other τεχνῖται even in matters pertaining to the latter's fields of expertise. Indeed, 'there is no subject on which the rhetorician could not speak more persuasively than a member of any other profession whatsoever, before a multitude.'³⁰ For these reasons, rhetoric may be said to 'comprise in itself practically all powers at once'³¹, and to be 'a cause not merely of freedom to mankind at large, but also of dominion to single persons in their

²⁸ *Go* 500c1–4

²⁹ *Go* 452d1ff.

³⁰ *Go* 456c4–6

³¹ *Go* 456a7–8

several cities.’³² Rhetors are the most powerful citizens in the state and the chief consultants in all public matters. If they wish, they can have doctors as their slaves, along with hunters, soldiers, politicians and everyone capable of advancing their personal interests.

But the fact remains that rhetors, inasmuch as they lack the actual knowledge provided by the different τέχνηαι, are ignorant of the matters they persuade others about. Socrates’ first charge against Gorgias’ definition of rhetoric consists in highlighting the fundamental difference between knowing or having learnt that something is true (or false) and believing or having been told so.³³ Since rhetors are not experts in the proper sense of the word, the kind of persuasion they produce generates belief without knowledge, and is thus unable to instruct others about true and false beliefs, right and wrong decisions, aims to be pursued and to be avoided, etc. Furthermore, since rhetoric is really only able to produce the appearance of knowledge, rhetors can only persuade those who are as ignorant as themselves. Doctors and hunters may well be talked by rhetors into putting their expertise at their disposal, but rhetors cannot persuade doctors or hunters about medical or hunting matters. For the beliefs generated by rhetoric are not as persuasive as those derived from actual knowledge.

What, then, is the origin of rhetoric’s power? Although rhetors do not know the things doctors know about medicine, nor the things hunters know about hunting, they do possess some kind of knowledge. But instead of a proper τέχνη, this knowledge amounts to a mere ‘habit’ (ἐμπειρία),³⁴ whose real aim is to produce pleasure and gratification. Socrates proceeds to illustrate this claim by selecting different branches of knowledge and dividing them into two separate lists: the first includes the arts related to the welfare of the body (gymnastic and medicine) and the soul (justice and legislation); the second includes different activities also related to the body (cookery and self-adornment) and the soul (rhetoric and sophistry).³⁵ The main difference between these two sets lies in the final aim of each activity: whereas gymnastic, medicine, justice and legislation aim to determine what is best for the body and the soul, cookery, self-adornment, rhetoric and sophistry aim to please the body and to flatter the soul.

The doctor is in possession of a τέχνη, while the cook cultivates a mere habit. Yet both medicine and cookery claim to give the body what is best (τὸ βέλτιστον). The disagreement stems, therefore, from a different understanding of what the best

³² *Go* 452d6–8

³³ See *Go* 454c.

³⁴ Cf. *Phdr* 270b5–6, where Socrates draws a distinction between τέχνη and τριβὴ καὶ ἐμπειρία, i.e. practice and experience. In both cases, Plato is pointing out that the kind of knowledge conveyed by rhetoric is not epistemic, but empirical. Unlike doctors and other experts, rhetors are ignorant of the truths acquired through a serious study of medicine or hunting. Nonetheless, they do know, out of observation and practice, what other people usually consider true or false, good or bad, desirable or undesirable, etc., and this knowledge allows them to conquer their hearers and influence their opinions.

³⁵ See *Go* 464b–466a.

might be: whereas doctors consider the best to be ‘the healthiest’ (τὸ ὑγιέστατον), cooks regard the best as ‘the pleasantest’ (τὸ ἡδιστον); and while doctors base their opinions on their knowledge of anatomy and physiology, cooks rely on their knowledge of the human palate. In the case of cookery, the good and the pleasant coincide because the pleasure conveyed by well-cooked dishes is instantaneous. As for medicine, however, the good and the pleasant often contradict each other, and only coincide when the patient is healthy and the medical art no longer needed. The reasons for this disagreement are twofold. First, a state of disease is not always noticeable. One may lead the pleasantest of lives and suffer, at the same time, from the worst of diseases. Second, there is nothing pleasant about a bitter medicine or a painful therapy. They are chosen not for their own sake, but as a means of securing a final and greater good, namely the decrease of pain or the cure of one’s illness. However, precisely because the medical road towards the good is often long and unappealing, the healthiest options are neglected in favour of more immediate pleasures.

The parallel between medicine and cookery is intended as an illustration of the parallel between justice and rhetoric. Since the rhetors’ eloquence enables them to steer the will of the crowd, their voice carries a lot of weight with the state. They are invited to give advice on important public matters and to help decide which laws are just and unjust or which courses of action should be followed. Like cooks, however, rhetors are not real experts, but ‘shrewd, gallant’ spirits with ‘a natural bent for clever dealing with mankind.’³⁶ They do not possess a method to distinguish justice from injustice, but only the ability to determine which decisions are likely to please the crowd. Just like cooks, they prescribe what is pleasant, but not necessarily what is best. And their decisions, however flattering, may come to harm the citizens and the state.

Pressed by Socrates, Gorgias ends up conceding that good rhetors, if they are to be more than mere demagogues, must learn what is just and act accordingly. But Polus rejects this idea. For him, the fact that most rhetors are not concerned with justice, but merely with their own self-interest, does not restrict the power of rhetoric. On the contrary: if rhetors are not bound by moral or political scruples, their versatility is even greater. They are free to talk anyone into saying or doing anything and their power is thus potentially unlimited.

At this point, the dialogue enters its true subject matter. The debate between Polus and Socrates turns into a debate about power, its nature and the kind of freedom it affords. More specifically, it turns to the issue of whether power is *intrinsically* advantageous or must be *rendered so* through the acquisition of knowledge. Polus likens the rhetor’s power to that of a despot. Although the deeds of despots are not always just, their authority compensates for their iniquities and protects them from harm. What is more, Polus argues that their limited regard for justice echoes an attitude that is instinctively, albeit secretly shared by most of their subjects. Even though most citizens condemn crime and injustice, they do so out of respect for

³⁶ *Go* 463a7–8

public order and fear of the law. But they would all be glad, if given the chance, to swap places with a powerful and unjust ruler. Indeed, who would not lay aside his or her moral qualms in exchange for ‘the liberty of doing anything one thinks fit in one’s city’, including ‘putting people to death and expelling them and doing everything at one’s own discretion’?³⁷ If the fear of punishment were eliminated, and wrongdoing were rewarded with success and impunity, most people would choose power over justice and sacrifice rectitude to the actual advantages of a privileged life.

For Socrates, on the contrary, rhetors and despots are not only powerless, but in fact the least powerful people in the whole city (466b-e). And to the dismay of Polus and Callicles, this claim is grounded in an even more daring one: to do wrong is the greatest of evils—greater still than to suffer wrong, however great one’s suffering might be (469b-472e). Rhetors and despots, insofar as they do wrong, are not to be envied, but to be pitied. And their alleged freedom is no freedom at all.

These paradoxical claims might be read as a mere defence of justice against Polus’ Machiavellian standpoint. While Socrates sides with goodness and rectitude, Polus sides with evil and iniquity; while the former is a truly honest man, the latter is an opportunist. But this dichotomic reading fails to appreciate the true depth of Plato’s argument. Instead of a mere contest between good and evil, or justice and injustice, the debate between Socrates and Polus is a debate about knowledge—about the extent to which the things one knows and ignores is brought to bear in the definition of power, freedom and happiness.

According to Socrates, what unites rhetors, despots and ‘the many’ is their incessant search for τὸ βέλτιστον, i.e. *the best there is*: every human being, whether rich or poor, powerful or miserable, eloquent or tongue-tied, wants to lead the best possible life. But although all of us choose and act in view of what is best, most of us lack a solid criterion whereby to determine what it actually is. The reason for this, Socrates argues, lies in the difference between the things we *wish* and the things we *know*. Since we all wish what is best, we all do what we *think* is best; but since what we think is best is not necessarily so, we do not always do what we wish. And herein lies the contradiction underlying rhetoric’s and despotism’s supposedly unlimited power. For Polus, there is no doubt that everyone would welcome ‘the liberty of doing anything one thinks fit’. In his eyes, such a power is intrinsically good, and no one in his or her right mind would fail to realize it. For Socrates, such a power is the result of a determinate *thesis* about the greatest good, which requires demonstration. The power hailed by Polus can only be shown to be beneficial if it is shown to be good, and not merely claimed to be so.³⁸

³⁷ *Go* 469c5–7. See also *Re* 344b–d and 359b–360d.

³⁸ In Hegel’s words, the argument consists in showing that ‘this good, which has by me to be esteemed as a substantial end, must be known by me; with this the infinite subjectivity, the freedom of self-consciousness in Socrates breaks out.’ (LHP, vol. 1, 385f. / HW 18, 442)

To grant Socrates' point, one needs simply to return to the comparison between doctors and cooks. When a sick person takes a bitter medicine, he or she does something unpleasant in view of something pleasanter, and thus better. But when a diabetic eats too much sugar, he or she does something that is immediately pleasant, but might turn out to be unpleasant, and thus worse. In the case of rhetors and despots, 'the liberty of doing anything one thinks fit' might turn out to be harmful for a number of reasons. Confident in their own power, despots might overestimate the extent to which they are immune to retaliation and end up being harmed by those they have harmed. Alternatively, they might be free of danger but become prisoners of the need to reassert their authority and protect themselves from rivals and conspirators.³⁹

But Socrates is by no means a consequentialist. He does not believe despotism to be evil merely because the despot is not immune to being harmed. He claims, rather, that doing wrong is *itself* the greatest of evils—even greater than being harmed. And, moreover, he argues it is impossible for an unjust person to be happy, regardless of whether he or she is punished or comes to harm. Wrongdoers, *because they are unjust*, are 'wretched anyhow.'⁴⁰

To prove his point, Socrates starts by having Polus admit that doing wrong is 'uglier' or 'fouler' (αἰσχρόν) than suffering it. This admission sets the ground for the demonstration of his previous and more controversial claim that 'doing wrong is worse than suffering wrong, and even worse when the wrongdoer is left unpunished.' As regards the first half of this statement, Socrates' argument can be divided into three main steps: he argues a) that a 'fair' thing (καλόν) is usually deemed so because it is either pleasant and/or beneficial; b) that doing wrong can therefore only be deemed 'fouler' than suffering wrong if it is either less pleasant and/or less beneficial than suffering wrong; and c) that doing wrong, since it is not less pleasant than suffering it, must be less beneficial—which leads to the conclusion that doing wrong is indeed worse than suffering it. As regards the second half of Socrates' claim, his argument can also be divided into three main steps: he argues a) that a fair action is usually deemed so because it is either pleasant and/or beneficial; b) that punishing someone justly can therefore only be deemed fair if it is either pleasant and/or beneficial; and c) that a just punishment, since it is not pleasant, must be beneficial—which leads to the conclusion that doing wrong is indeed worse when the wrongdoer is left unpunished, *quod erat demonstrandum*.

³⁹This theme is further developed in the *Republic*'s ninth book, where despots are also said to live in constant fear of those they have wronged. See *Re* 578bff.

⁴⁰*Go* 472e5. This point is highly ambiguous. By arguing that doing wrong is worse than suffering it, and even more so when the wrongdoer is left unpunished, Socrates is doing more than simply denouncing the groundlessness and/or unintelligibility of Polus' standpoint. In contrast with the negative procedure that usually guides his criticism of δοξάζειν, he is advancing two positive claims about justice and wretchedness—which is at best atypical for someone whose usual concern is to suggest that no categorical claim is ever truly substantiated. Since the *Gorgias* fails to offer, up until the very end, an explicit justification of Socrates' claims, there seems to be no way of guaranteeing that they are not themselves doxastic, and hence as open to criticism as Polus'.

This demonstration prompts the intervention of Callicles, which opens the third and last section of the dialogue. Like Polus, Callicles starts by highlighting the counter-intuitive nature of Socrates' claims: if every wrongdoer is indeed wretched, and even more so if he or she escapes punishment, then 'the life of us human beings [must] have been turned upside down' and we must be doing 'quite the opposite, it seems, of what we ought to do'.⁴¹ Unlike Polus, however, Callicles is not willing to accept Socrates' association between fairness and goodness (or foulness and evil). For Callicles, Polus was not able to dodge Socrates' charges because he was not as bold as he should have been. Instead of minding about fairness and foulness, he should have dared to look for the truth about justice and injustice beyond the realm of social conventions.

Callicles' views on the merits and inconveniences of a just life echo most of the key points addressed by Socrates, Thrasymachus and Glaucon in the *Republic's* first and second books. At stake in both dialogues is whether justice is desirable in itself, as Socrates argues, or whether it is practised out of necessity, but against one's innermost will, as argued by Polus, Callicles and Thrasymachus. Just as Polus had maintained that no one in his or her right mind would choose a life of virtue over the life of a despot, Glaucon argues that for most people justice 'belongs to the toilsome class of things that must be practised for the sake of rewards and repute due to opinion, but that in itself is to be shunned as an affliction.'⁴² If humans created laws that favour justice over injustice, it was not because the former is superior to the latter, but due to 'a compromise between the best, which is to do wrong with impunity, and the worst, which is to be wronged and be impotent to get one's revenge.'⁴³ The respect for justice is thus artificially ensured by a 'social compact', devised by the weak and impotent to protect themselves against the strong:

When men do wrong and are wronged by one another and taste of both, those who lack the power to avoid the one and take the other determine that it is for their profit to make a compact with one another neither to commit nor to suffer injustice.⁴⁴

The aim of this compact is to curb human beings' natural appetite for power through an artificially imposed respect for justice and equality. For Callicles, Thrasymachus and Glaucon, this idea is at the origin of the distinction between what is true or just *according to law*, or *custom* (κατὰ νόμον), and what is true or just *according to nature* (κατὰ φύσιν).⁴⁵ Although a just life is socially commendable and practically necessary, lest one be marginalized or punished, it is by no means natural. For

⁴¹ *Go* 481c3–4

⁴² *Re* 358a4–6

⁴³ *Re* 359a5–7

⁴⁴ *Re* 358e5–359a2

⁴⁵ See *Go* 482e–484c and *Re* 359c, but also *Pro* 337d, *Phdr* 272d, *La* 715a and 890a. Though the debates in the *Gorgias* and the *Republic* are very similar, they also differ in significant ways. For a comparison, see Barney, 'Callicles and Thrasymachus'; Broze, 'Calliclès et Thrasymaque'.

nature is by definition indifferent to human conventions, guided only by the immediate laws of life, growth and strength.⁴⁶ There is thus a fundamental incompatibility between what is naturally good or evil and what is morally or legally so:

By nature everything is fouler than is more evil, such as suffering wrong: doing it is fouler only by convention. Indeed this endurance of wrong done is not a man's part at all, but a poor slave's, for whom it is better to be dead than alive.⁴⁷

Callicles' eulogy to natural justice evokes once again the theme of mastery and slavery. On his view, these two notions do not correspond merely to social or political roles, but to natural categories, which may or may not be actualized in the social and political realm. 'Natural' masters, on the one hand, are slaves to no one. They are free to do what they please whenever they please, to indulge their every desire without fearing any requital. This, of course, is not allowed to the many, but only to the 'better type of mankind',⁴⁸ to whom 'luxury and licentiousness and liberty', since they 'have the support of force, are virtue and happiness.'⁴⁹ On the other hand, the lot of 'natural' slaves includes all of those who lack the strength to enjoy such freedom. But since these amount to the vast majority of human beings, the 'natural' opposition between masters and slaves is overturned and replaced by the 'social compact' mentioned above: most people, in order to disguise their own powerlessness, decry luxury and liberty as shameful and unjust, thereby restraining other people's power. 'Being unable themselves to procure the achievement of their pleasures they praise temperance and justice by reason of their own unmanliness',⁵⁰ thus 'enslaving the better type of mankind'.⁵¹

*Fit in dominatu servitus, in servitute dominatus.*⁵² Although sovereignty and slavery amount to natural categories, determined by one's natural endowments, the 'unnatural covenants of mankind' have led masters and slaves to switch places. Opposed to this role reversal, Callicles prophesies the elimination of social conventions and the restoration of the natural order:

⁴⁶The translation of νόμος as 'custom', 'convention', etc., may be misleading. These words usually refer to a set of codes and practices accepted by the members of a community. However, when obeying these rules, they are often aware that they are artificial, i.e. mere customs or conventions. What Callicles and Thrasymachus have in mind, however, is a stronger form of obedience. They are referring to νομοί that are not lived as νόμμοι, but as φυσικοί. And herein lies the force of their argument: they are denouncing what appears natural as an ideological fabrication, designed to benefit only a few members of the community.

⁴⁷Go 483a7–b2. The distinction between νόμος and φύσις is a central theme of the sophistical tradition. See Menzel, *Kallikles*, 1–75; Guthrie, *A History of Greek Philosophy*, 55–134.

⁴⁸Go 492a6–7

⁴⁹Go 492c4–6

⁵⁰Go 492a7–b1

⁵¹Go 492a6–7

⁵²Cicero, *Orations*, 30.

When some man arises with a nature of sufficient force, he shakes off all that we have taught him, bursts his bonds, and breaks free; he tramples underfoot our codes and juggleries, our charms and 'laws', which are all against nature; our slave rises in revolt and shows himself our master, and there dawns the full light of natural justice.⁵³

Like Callicles, Socrates also speaks of mastery and slavery, but in very different terms. On his view, freedom is not primarily based on power or strength, but on knowledge. To refute Callicles' position, he subjects his conception of freedom to his usual zooming method. He starts by focusing on Callicles' distinction between a 'better' and a 'baser part of mankind': when asked what he means by superior and inferior sorts of people, Callicles offers a series of different, but equally inconsistent definitions. At 488d he claims that some human beings are superior to others due to their physical strength; at 489e he associates superiority with wisdom (or soundness of mind, *φρονιμότης*); at 491b with courage and manhood; and at 491e with fullness and gratification (*πλεονεξία*).⁵⁴ In this last definition, the essence of Callicles' standpoint is finally brought out. His idea of freedom comes down to an extreme form of hedonism, where power is measured by the ability to 'let [one's] desires be as strong as possible, and not chasten them', and 'to minister to them when they are at their height', and 'to satisfy each appetite with what it desires.'⁵⁵

This definition leads back to Polus' standpoint and to the distinction between the good and the pleasant. The dialogue's final debate amounts, therefore, to a development of Plato's initial argument concerning the relationship between freedom and knowledge: since a) what is pleasant or desirable is always pursued in view of what is best, and not otherwise, ergo b) the ability to do what one wills is only liberating insofar as it is actually beneficial. In other words, true freedom rests on the ability to distinguish good or beneficial pleasures from bad or harmful ones. And such ability presupposes a specific knowledge, provided by a specific *τέχνη*. For Socrates, the name of this *τέχνη* is none other than philosophy, 'a certain practice or preparation' for the acquisition of the good.⁵⁶ Following this new definition, the initial dispute between Socrates and Gorgias concerning the merits of rhetoric is recast under a new light: whereas the rhetor's art consists in the procurement of pleasure or gratification, naïvely understood as naturally or immediately beneficial, the philosopher's role is to expose the ambivalent nature of pleasure and to distinguish good pleasures from false or illusory ones; moreover, whereas Gorgias had defined rhetoric as the finest of all arts, 'comprising in itself . . . all powers at once', this epithet is now shown to belong by right to philosophy. Unlike the freedom provided by medicine, hunting or any other particular *τέχνη*, the freedom pursued by philosophy is not limited to a specific domain of existence. By attempting to define what is best, the philosopher's aim is to lay out the very ground on which freedom of any kind can flourish.

⁵³ *Go* 484a2–b1

⁵⁴ The word *πλεονεξία* only comes up later, at 508a7, but Plato uses it to sum up this passage and its subsequent developments.

⁵⁵ *Go* 491e8–492a2

⁵⁶ *Go* 500d8–9

We are now in a better position to understand the Platonic reversal of Callicles' master-slave dialectic. When distinguishing between masters and slaves, or free and imprisoned human beings, Plato does not refer merely to the political meaning of these words, as we often do, or to their biological meaning, as Callicles does. He is suggesting, rather, that *all of us*, insofar as we lack the actual knowledge of what is best, are naturally enslaved.⁵⁷ Although convinced we are free to choose what we want, we may in fact be doomed to choose what we do *not* want—or what is not, after all, in our best interest. We find here the essence of the contradiction that will reappear in the allegory of the cave and in other Platonic texts: the greatest obstacle to human freedom does not lie outside human consciousness, but within it, in the very idea of freedom. The chains that bind us are the chains of *δοξάζειν*, and the only chance of release is a philosophical reform of human cognition.

Plato's conception of freedom entails a radical transformation of our usual cognitive perspective, bound to affect all our ideas and beliefs. The universal scope of this transformation, suggested throughout the *Gorgias*, is also highlighted in the *Theaetetus*, following Plato's critique of rhetors and litigators. As we have seen, when describing the life of philosophers, Socrates stresses their lack of common sense and their ignorance in public and political matters: whenever a philosopher meets with other people, or is made to speak in court or elsewhere, he soon becomes 'a laughing-stock not only to Thracian girls but to the multitude in general, for he falls into pits and all sorts of perplexities through inexperience.'⁵⁸ Yet his 'inexperience' refers to the things deemed important by the 'multitude in general', while his 'perplexity' is directed at the things the multitude claims to know and understand. If these things are neither inherently important nor inherently clear, but merely assumed to be so, inexperience and perplexity may end up changing sides. If so, the alleged superiority of 'reasonable people' disappears and the alleged enslavement of philosophers emerges as the only real path to freedom:

When, my friend, [the philosopher] draws a man upwards and the other is willing to rise with him . . . to the investigation of . . . human happiness and wretchedness in general, to see what the nature of each is and in what way man is naturally fitted to gain the one and escape the other—when that man of small and sharp and pettifogging mind is compelled in his turn to give an account of all these things, then the tables are turned; dizzy by the new experience of hanging at such a height, he gazes downward from the air in dismay and perplexity; he stammers and becomes ridiculous, not in the eyes of Thracian girls or other uneducated persons, . . . but in those of all men who have been brought up as free men, not as slaves.⁵⁹

⁵⁷ Ironically, Plato himself was once sold into slavery—actual slavery—and eventually released for twenty minas. According to Diogenes Laertius, he was freed when someone told his owner he was a philosopher—the implication being that philosophers are useless creatures, with a very low market value. See Riginos, *Platonica*, 86–92.

⁵⁸ *Th* 174c3–5

⁵⁹ *Th* 175b9–d7

Chapter 4

Chains and Shadows



Abstract In Plato's writings, men and women are often compared to prisoners or slaves. This chapter explores this theme by returning to its canonical formulation, found in the so-called allegory of the cave. By examining the allegory's different elements, I show that Plato's focus on freedom and imprisonment is not merely a metaphorical or literary device, but a new and revealing approach to the debate about freedom's cognitive dimension.

Section 4.1 offers a detailed analysis of the allegory's structure and starting point, with a special emphasis on Plato's explicit comparison between the cavemen's cognitive imprisonment and the one endured by most human beings. Section 4.2 examines the liberation of one of the prisoners and the ascent towards the cave's entrance. Drawing on Socrates' description, I discuss the different ways in which this movement reflects Plato's project of reforming and liberating human cognition. Finally, in Sect. 4.2, I argue that the allegory, apart from its epistemological significance, also speaks to the possibility of a radical ontological transformation. Besides the shadows projected on the wall, the prisoners themselves lead a shadow-like existence, whose exact characterization is crucial to understanding Plato's main argument.

Keywords Plato · Socrates · *Republic* · Allegory of the Cave · Prisoners · Chains · Recollection

In the dialogues, human beings are often said to be imprisoned or enslaved. In the two previous chapters, we have seen some of the central ways in which Plato articulates this idea: first, regarding the basic structure of human cognition, I stressed his insistence on our usual tendency for self-deception, illustrated by the soul's confinement to the body; second, regarding his definition of philosophy and rhetoric, I discussed the different meanings ascribed by Socrates and his adversaries to the notions of power, freedom, mastery and slavery. In this chapter, I explore this Platonic theme by returning to its canonical formulation: the allegory of the cave.

This well-known image is summed up by Socrates in a few brief sentences. In the bottom of a dark, subterranean cave, a group of prisoners is forced to face a stone

wall. Their backs are turned to the entrance of the cave, located at the end of a long upward path, to a fire burning somewhere along this path and to a low wall resembling a screen used in marionette performances, placed between the fire and the prisoners. Various objects are held above the screen and their shadows are projected on the wall at the bottom of the cave. The prisoners watch the movement of the shadows and hear the echoes of voices but are unaware of the performance taking place behind them.

Before focusing on the different elements of this description, it is important to consider Plato's overall strategy. As we have started seeing earlier, the allegory of the cave is a complex image, which can be looked at from very different angles. At the beginning of the allegory, two different standpoints are at stake. The first, and most basic one, is the prisoners' outlook on reality, limited to the shadows projected on the wall. Since the prisoners' bodies are bound by chains, and have always been so, their state of imprisonment is not lived or acknowledged as such. Moreover, since the prisoners' necks are also bound by chains, which force them to look ahead, the wall and the shadows are all they know. From where they stand, the shadows are not shadows, but real things, and the limits of the wall coincide with the limits of the world.

The second standpoint at stake in the allegory is our own perspective as readers and interpreters. Unlike the prisoners, we have access to a wider portion of reality. We know that the wall is not the whole world and that the images on it are mere shadows of real objects, physically placed elsewhere. The prisoners' life amounts, therefore, to a limited version of our own life. Everything changes, however, when Socrates claims that the prisoners are *like us*, and that their situation is *our situation*.¹ This declaration introduces a turning point in the narrative, shifting the focus from the prisoners' cognitive limitations to our own cognitive limitations. The allegory started out as the illustration of a cognitive scale in which we occupy the highest position and the prisoners an inferior one, but this turning point introduces the possibility of a wider cognitive scale, in which our perspective and the prisoners' are both downgraded versions of a higher cognitive perspective. Plato's allegorical procedure amounts to what I called earlier, following Fränkel, a Heraclitean thought pattern²: Socrates resorts to a comparison whose terms we *can* understand as a means of introducing a second comparison, of which only one term is familiar to us. Starting from the analogy between our own life, usually supported by an immediate belief in freedom and clairvoyance, and that of a group of prisoners who are equally convinced of their freedom and clairvoyance, he establishes a new analogy, in which we are accorded the prisoners' former position and our own position is shifted forward: just like the prisoners are fettered and fail to notice it, we too may be fettered and fail to notice it; and just as they inhabit a cave which they mistake for the whole world, we too may be trapped inside a cave which we mistake

¹Re 515a5

²See Sect. 4.2.

for the whole world. To understand Plato's suggestion, we must project the difference between the prisoners and ourselves into an unknown cognitive terrain and posit a new human being ahead of us—one who is somehow freer and more knowledgeable than we are.³

4.1 Inside the Cave

Plato's initial description refers not to a simple imprisonment, but to a double one. Inside the cave, the prisoners are subjected both to a *physical* confinement, due to the chains that bind them, and to an *optical* confinement, due to having to see shadows instead of real objects. Both of these restrictions reflect specific aspects of human cognition, which must be briefly considered.

The first form of imprisonment speaks to a paradox that is highlighted in different ways in Plato's dialogues. As already noted, since the shackles fastened round the prisoners' legs and necks are in place since early childhood (ἐκ παίδων), their captivity is the only kind of freedom they know. Unaware of their own bodies, or the other prisoners, or the surrounding space, they are incapable of grasping their real situation or appreciating the real extent of their freedom. The very notions of freedom and captivity, along with every other idea or belief, are conveyed by the shadows on the wall. But although the prisoners are liable to spend their entire lives in this state of delusion, their situation is nonetheless artificial and maintained by force. If they were depicted in such a way that their anatomy would somehow prevent them from moving their heads and looking in different directions, their

³The difference between the allegory's *illustrans*—a community of prisoners held inside a cave—and the allegory's *illustratum*—the prisoners as symbols of ourselves and our own life—can be further clarified by comparing the structure of the allegory of the cave and that of the so-called allegory of the divided line, featured in the *Republic*'s sixth book (esp. 509d–511e). As is generally known, the line imagined by Socrates to illustrate human cognition's different hierarchical stages is divided into two main sections, referring to the visible world (τὸ τοῦ ὁραμένου) and the invisible world (τὸ τοῦ νοουμένου). These sections are each divided into two subsections. The four resulting subsections refer, in ascending order, to the cognition of images or shadows of real things (εἰκόνες), the cognition of real objects and living beings (ὁρατά), the cognition of abstract or intellectual notions and the cognition of true ideas. Insofar as the prisoners inside the cave are forced to see shadows of real things, their cognitive perspective is initially situated in the line's first subsection. And since the allegory documents, as we will see below, the transition from this world of shadows to the real world, located outside the cave, the prisoners' final destination is the line's second subsection. But this correspondence refers only to the allegory's *illustrans*, and not yet to its *illustratum*. Since *we*, unlike the prisoners, already inhabit the real world, *our* starting point is the line's second subsection and *our* destination is the line's second main section. In *our* case, the allegory documents the possibility of adopting a higher cognitive standard, which reveals what we usually take to be real and intelligible as unreal and unintelligible. The transition from the third to the fourth subsection corresponds, therefore, to the basic experience enacted in most of Plato's dialogues: starting from our usual cognitive perspective, Socrates calls attention to its inner contradictions and to the need for a new and more solid outlook on reality.

cognitive limitations would be part of their original identity. In that case, there would be no real tension between what they see and what lies behind them, because their limited visual scope would coincide with their natural visual horizon. But Plato's prisoners, just like us, are *potentially* able to look in different directions and see different things. Their captivity is not natural, but imposed: it prevents them from achieving their full cognitive potential and reduces them to a lesser or diminished version of themselves. In other words, the knowledge they are able to attain exceeds the knowledge they actually possess, and this idea leads back to the paradox we have encountered when discussing the Platonic contrast between body and soul. As suggested by the *Phaedrus*' flying chariot, there is an original disproportion between the soul's natural cognitive vocation and its actual cognitive condition—a yearning for truth and clarity that is usually quenched by a vast web of cognitive assumptions.

Yet the captivity described by Socrates is not due only to the chains that bind the prisoners, forcing them to see shadows. It is also caused by the shadows themselves, and by their peculiar mode of identity. In the allegory, as in real life, shadows amount to imperfect representations of real things, produced by objects interposed between a luminous source and a given surface. Depending on the relative positions of the objects, the luminous source, the surface of projection and the observer, shadows can have different degrees of sharpness—ranging from clear-cut black projections to diffuse greyish stains—and different degrees of distortion—from familiar silhouettes to unrecognizable shapes. In Plato's allegory, however, the most important thing about shadows is their *borrowed* or *second-hand* nature. They amount to a particular instantiation of what Plato often refers to simply as 'images' or 'copies' of real things (εἰκόνες, εἰδωλα)—as when he argues, for example, that our usual account of reality is only an *image* of reality, or that our usual ideas are only *images* of true ideas.

Unlike real objects, shadows are not direct sources of meaning. When considered in isolation, disconnected from the objects they refer to, they cease being shadows and become abstract shapes. Their definition is intrinsically tied to the object that produces them, whether or not that object is known, and their identity is thus inherently transitive. Shadows are visual signposts, as it were, or visual reminders of real objects. They are what they are by pointing to something else. Whenever we see a shadow, we do not see a mere shape, but the image of a specific object or a specific person. And even if we do not see or cannot identify that object or person, we nonetheless assume that it must exist somewhere. In short, to identify an unknown shadow is not to say what it is *in itself*, but what it is *a shadow of*.

In Plato's cave, we readers know that the shadows are mere shadows and that their true identity must be looked for behind the prisoners' backs. From where we stand, the whole situation comes down to a simple lack of correspondence: if the prisoners were allowed to turn round, they would find out what we already know, and would link the shadows on the wall to the objects above the screen. From the prisoners' perspective, however, things are very different. Since they have never seen anything other than the shadows on the wall, they are not aware of their

shadow-like quality. In their eyes, the shadows amount to real, self-identical objects, whose meaning must be looked for in themselves. The world perceived by the prisoners is radically different from the real world, but since the very standard with which the prisoners measure what is real is provided by shadows, this difference goes unnoticed.

If the prisoners' perspective were presented in purely conjectural terms, as the peculiar outlook of a peculiar sort of creatures, its interest would be merely theoretical. But with Plato's Heraclitean move, readers are invited to abandon their omniscient standpoint, to step into the prisoners' shoes and to take their predicament more seriously. By suggesting that they are like us, Plato is implying that our own world—the world we see, touch and hear every day—is entirely made up of shadows. As extraordinary as this suggestion might seem, its scope and gravity must not be minimized: Plato is implying that none of the objects we see or come into contact with are in any way real or self-evident. They are all mere shadows, wrongly perceived as real objects.

But this is not all. Plato is implying, furthermore, that we are surrounded by shadows produced by objects *we have never seen or experienced*. And this suggestion is even less intuitive, as it challenges the very idea of shadow we are familiar with. Since shadows are what they are by referring to something else, their definition presupposes the existence of an object, whatever it may be. Usually, then, *objects precede shadows*: they grant them their meaning and purpose, the backdrop against which shadows can be thought of as such. Yet by reducing our standpoint to that of the prisoners, Plato is turning this basic model upside down. In human life, he argues, *shadows precede objects*. Our account of reality is primarily based on shadows, not objects, which means not only a) that all the things we conceive and experience are mere signposts pointing *somewhere else*, but also b) that all the things we conceive and experience amount to *empty* signposts, or to signposts pointing in an utterly unknown direction.

But how exactly does Plato's suggestion relate to everyday life? And why, if at all, should it worry us? To answer these questions, let us abandon for a moment the Platonic cave and turn our attention to two Platonic notions brought up in other dialogues, but no less present, albeit only implicitly, in the *Republic*'s allegory. The first one is the notion of μέθεξις, i.e. 'inherence' or 'participation'. Its meaning can be briefly clarified by turning to the *Phaedo* and to the widely quoted passage in which Socrates discusses the definition of beauty (or fairness, or nobility: τὸ καλόν):

I think that if anything is beautiful besides beauty itself [αὐτὸ τὸ καλόν] it is beautiful for no other reason than because it *partakes* [μετέχει] of beauty itself; *and this applies to everything*.⁴

Shortly afterwards, Socrates adds:

⁴ *Phd* 100c4–6. Emphases added.

If anyone tells me that what makes a thing beautiful is its lovely colour, or its shape or anything else of the sort, I let all that go, for all those things confuse me, and I hold simply and plainly and perhaps foolishly to this, that nothing else makes it beautiful but the presence or communion (call it which you please) of beauty itself . . . I do insist that beautiful things are made beautiful by beauty. For I think this is the safest answer I can give to myself or to others, and if I cleave fast to this, I think I shall never be overthrown.⁵

Whenever we ask, as Socrates often does, what it means for something to be beautiful, or coloured, or solid, etc., we are inevitably led back to a set of general categories, or εἶδη. Something is beautiful because it *participates* or *inheres* in the notion of beauty, just as something is white because it *partakes* of whiteness, and so on. When we say ‘white’, we ‘cannot help thinking’ of ‘whiteness’. And according to Socrates, to say that something is white because it partakes of whiteness is ‘the safest answer we can give’.

At this point, two brief clarifications are in order. Firstly, although many authors translate Plato’s αὐτὸ τὸ x’ as ‘absolute x’, this solution is ambiguous and tends to misrepresent Plato’s argument. In this and similar passages, Socrates is not qualifying beauty in terms of its quantity or intensity, as though the expression αὐτὸ τὸ καλόν were meant to single out a ‘concentrated’ or ‘saturated’ form of beauty, as opposed to a ‘faded’ or ‘diluted’ one. Such a reading usually leads to the simple dualism found in most accounts of Plato’s so-called Theory of Ideas, i.e. to the simple contrast between human beings’ earthly existence and a supernatural existence, where concepts are somehow livelier or purer than the ones employed in real life. But what is at stake in this and similar passages is not a *quantitative* distinction between different degrees of beauty, or fairness, or whiteness, or anything else. If that were so, we would already be in possession of an adequate conception of beauty, for example, and would simply have to look for a clearer or livelier one. Plato’s argument is rather about the *qualitative* opposition between *true* beauty, or beauty *in itself*, and a mere *version* of beauty, wrongly perceived as the real thing.⁶ The point is that our usual conception of beauty does not coincide with beauty itself because it is only a mere δόξα, or a mere image, or a mere shadow of beauty. Although we speak of beauty as though we knew what it means, we are dealing with a *transitive* or *second-hand* determination, whose meaning is not readily available.

The second clarification relates to the scope of Plato’s argument. In the *Phaedo* and elsewhere, Socrates’ discussion of μέθεξις is centred on elementary concepts such as beauty, goodness, unity or difference.⁷ These concepts are the most basic

⁵ *Phd* 100c10–e1. The universality of this principle is also highlighted in *Phi* 15d–e: ‘We say that one and many [viz. abstract notions and their particular instantiations] are identified by reason, and always, both now and in the past, circulate everywhere in every thought that is uttered. This is no new thing and will never cease; it is . . . a quality within us which will never die or grow old, and which belongs to reason itself as such.’

⁶ Plato’s ideal realm ‘is not beyond reality, in heaven, in another place, but is the real world; . . . The ideal is not metaphysical, but reality itself brought closer.’ (LHP, vol. 2, 29 / HW 19, 39. Translation modified.)

⁷ See *Phd* 100cff.; *So* 256b, 259a; or *Pa* passim.

ingredients of human cognition, and therefore play a privileged role in the debates. But this does not mean that the rule of μέθεξις only affects a small part of our cognitive horizon. Rather, as Socrates points out, it ‘applies to everything’, from the most general to the most specific human beliefs. It also applies, indirectly, to every combination of beliefs, and indeed to the whole web of assumptions underlying our usual outlook on reality. Every complex mental representation partakes of simpler conceptual units, which partake in turn of a series of concepts or ideas ‘in themselves’. The whole edifice of human reason is grounded in a thoroughly transitive web of beliefs, intrinsically related to concepts or ideas ‘in themselves’.⁸

According to Socrates, ‘there is no other way by which anything can come into existence than by participating in the proper essence of each thing in which it participates.’⁹ Therefore, to say that something is beautiful, fair, or white because it partakes of beauty, fairness, or whiteness is ‘the safest answer we can give’. But it is by no means a good answer. For when asked about what beauty or whiteness *are*, we can only provide examples of beautiful or white things. While every specific belief leads back to its εἶδος, or to a more or less complex combination of εἶδη, the latter can only be defined in relative terms, as *that in which each belief inheres or participates*. And this circularity—the circularity of μετέχειν, or μέθεξις—is precisely what is highlighted in Plato’s cave. The prisoners can only see shadows because they can only know things as particular instantiations of real objects. But these objects are themselves posited through their shadows, that is, through their particular instantiations. The relationship between shadows and real objects, or between beliefs and actual knowledge, is not disjunctive, but dialectical. Both terms are essential moments of the relationship and neither can be thought of without the other. Yet precisely because of their intrinsically transitive nature, neither shadows nor real objects provide a stable ground in which to anchor a true account of reality. Whether we cling to our beliefs or invoke the authority of a concept in itself, we are bound to deal with shadows, *and nothing else*.

Another Platonic notion that is directly, albeit inexplicitly linked to Socrates’ allegorical description, is the notion of ἀνάμνησις, or ‘recollection’, along with the well-known Platonic claim that ‘learning is nothing else than recollection.’¹⁰ In the *Phaedo*, when discussing the idea of equality, Socrates distinguishes between equality ‘in itself’ (αὐτὸ τὸ ἴσον) and the equality predicated of different things. Although the notion of equality is not usually considered in isolation, but in connection with concrete things—equal pieces of wood, for example, or equal pieces of stone—it is nonetheless irreducible to its particular instantiations. ‘Abstract’ equality is not the same as ‘concrete’ equality because equal pieces of wood or

⁸ Plato tends to speak of εἶδη in connection with basic or elementary ideas, as opposed to his use of δόξα, which is more generic. But this difference does not indicate a qualitative distinction between different kinds of beliefs. Inasmuch as human cognition is supported by a global web of ideas, combined in different ways, it is entirely *eidetic*; and inasmuch as this web is automatically assumed to valid and intelligible, it is entirely *doxastic*.

⁹ *Phd* 101c2–4

¹⁰ See notably *Phd* 72e–73b and *Me*, 80dff.

stone are never perfectly equal to each other: however similar, they can always be looked at from different angles, placed in different positions, lighted in different ways and perceived in accordance to different cognitive sensibilities, priorities or concerns. In short, although stones and pieces of wood may be considered equal, they ‘fall short of being like abstract equality.’¹¹ Or rather, as Plato also puts it, their equality *yearns or longs* (ὀρέγεται) to be like abstract equality but falls short of it.¹²

Our usual knowledge of equality is irreducible to the knowledge of equality conveyed by our encounters with equal things. Abstract equality is not a product of experience, but a condition thereof—not the result of an empirical observation, but the very standard whereby we observe and measure the equality of the things we encounter.

When anyone on seeing a thing thinks, ‘This thing that I see aims at being like some other thing that exists, but falls short and is unable to be like that thing’, he who thinks this must necessarily have previous knowledge of the thing which he says the other resembles but falls short of.¹³

Since our knowledge of equality transcends the realm of experience, it must have been acquired before we began to see, or hear, or use any of the senses. In other words, it must have been acquired before we were born. Simultaneously, however, since our knowledge of equality is not perfect or immediate, but reduced to the concrete instantiations of equality conveyed by equal things, it must have been forgotten at birth, or soon after we were born. These two conclusions are the two main principles of Plato’s doctrine of recollection, devised to account for the peculiar nature of our understanding not only of ‘the equal’, but also of ‘the greater and the less’, and of ‘all such abstractions.’¹⁴ In everyday life, we neither know what equality is nor ignore it completely: we learn it by recollecting what we once knew, but lost touch with. To learn is therefore to bring to mind a knowledge that is somehow already present within us, but must be recovered or reactivated.¹⁵ On the one hand, if we did not carry with us the idea of equality, we would not be able to recognize it in equal stones or pieces of wood; on the other hand, if we had not forgotten it, or lost touch with it, we would not be reduced to the imperfect versions of equality conveyed by equal stones and pieces of wood. As it is, we are trapped

¹¹ *Phd* 74d6–7

¹² *Phd* 75af

¹³ *Phd* 74d9–e2

¹⁴ *Phd* 75c9–10

¹⁵ By tracing our knowledge of the different εἶδη back to the soul’s disembodied existence, Plato uses a temporal image to describe what modern idealism would later call *a priori* knowledge. This parallel was highlighted by the neo-Kantian school (see Hartmann, *Platos Logik des Seins*, 180 f., or Natorp, *Platos Ideenlehre*, 41), but the clearest formulation of the distinction between the temporal and transcendental understandings of the concept of recollection can be found directly in Kant, in the classical definition of *a priori* knowledge offered in the *Critique of Pure Reason*: ‘As far as time is concerned, then, no cognition in us precedes experience, and with experience every cognition begins. But although all our experience commences *with* experience, yet it does not on that account all arise *from* experience.’ (B1. Emphases added.)

between knowledge and ignorance, divided between our *idea* of reality and our *experience* of it. Although equality, greatness and whiteness in themselves are irreducible to concrete equality, greatness and whiteness, they can only be known in a negative way, as *that which concrete concepts are recollections of*. The ideal and the empirical moments are intrinsically correlated and cannot be thought of in isolation.

Once again, this general description requires two brief clarifications. The first one relates to the tendency, already noted earlier, to interpret Plato's αὐτὸ τὸ ἴσον, or αὐτὸ τὸ ἔλαττον, or αὐτὸ τὸ καλόν, as quantitatively enhanced versions of familiar concepts, located in a 'super-world' of Ideas or Forms. This kind of reading is often coupled with a similar reading of Plato's doctrine of recollection, which tends to reduce the argument to a simple contrast between the soul's earthly existence and its spiritual or supernatural form. In this and similar interpretations, the *immanent* circularity of human cognition is explained away in a *transcendent* manner: like the reader's omniscient outlook on Plato's cave, the interpreters place before them both the natural and the supernatural realms and decree their incommunicability. In doing so, however, they fail to appreciate the intrinsically dialectical nature of recollection. They construe the object of the soul's recollection as an enhanced or intensified version of a regular object, placed beyond the confines of physical life, thereby avoiding the true implications of Plato's criticism.

This simplistic reading is also associated with an equally simplistic interpretation of the role played by mythology in the Platonic corpus. To be sure, the myths concerning the pre-existence and reincarnation of the soul are essential components of the dialogues. On several occasions, Socrates refers to εἶδη as notions learned in another life and to ἀνάμνησις as a recollection of what went on before our physical existence.¹⁶ But the meaning of these images tends to be reduced to one of two limited interpretations. The first consists in reading Plato's myths as literal descriptions, inspired by religious or mystical beliefs, and/or by the poor scientific resources available during his lifetime. This condescending approach reduces the mythological aspects of the dialogues to naïve cosmological descriptions, surpassed by modern scientific accounts of human cognition. The second main approach consists in reading Plato's myths as purely metaphorical or poetical descriptions, put down to Socrates' fertile imagination and separated from the more serious or properly philosophical sections of the dialogues. This reading, albeit more sophisticated than the previous one, also fails to appreciate the myths' intrinsic philosophical significance, treating them as secondary or downright superfluous elements.¹⁷

¹⁶See, for ex., *Phd* 76d-e and *Phdr* 249e-250c.

¹⁷While Leibniz rejects the 'error of pre-existence' in Plato's otherwise 'excellent' doctrine of anamnesis (*Die philosophischen Schriften*, vol. 4, 451f.), Natorp considers the 'whole mythical-mystical covering of the doctrine of recollection' a mere 'poetical addition' (*Platos Ideenlehre*, 35). More recently, against this line of interpretation, different scholars have highlighted the importance of the literary and dramatic elements of Plato's dialogues, including their mythical components. See, for example, Blondell, *The Play of Character in Plato's Dialogues*; Statkiewicz, *Rhapsody of Philosophy*; or Smith, 'The Ancient Quarrel and the Dream of Writing'.

In my view, the importance of mythology in Plato's dialogues can only be truly grasped by focusing not on the myths themselves (i.e. on what they tell us about eternal life, or metempsychosis, or the supra-celestial realm), but on the issue of human cognition (i.e. on what they tell us about ourselves and the way we cognize reality). Although the myths as such, prior to their allegorical function, are indeed fanciful tales, which may or may not be taken seriously, the myths as explanations designed to capture the essence of human cognition are as meaningful as the most serious philosophical theories. And this general principle is corroborated by Plato's usual philosophical method. Instead of speaking about migrating souls and the outer surface of heaven as of things existing in their own right, and afterwards trying to determine how they affect human life, Socrates usually does the opposite: having identified a given contradiction or inconsistency at the heart of human cognition, he then comes up with stories designed to bring it out in the clearest and liveliest ways.

Since Plato's myths amount to *ad hoc* explanations, their accuracy must be measured not by their degree of verisimilitude, but by the extent to which they succeed in capturing the peculiarity of our cognitive situation.¹⁸ Nevertheless, when suggesting that every particular determination presupposes a set of transcendental ideas *as though* they existed in some heavenly realm; or that human beings possess a latent knowledge of the various ideas *as though* they were learned in some previous life, Plato is not resorting to mere metaphors, in the conventional sense of the word. For a metaphor consists in replacing a direct meaning with an indirect one, which presupposes in turn that one has access to the direct meaning that is being replaced. This presupposition, however, is precisely what Plato keeps calling into question. The peculiar nature of our usual understanding of reality must be conveyed by means of images and myths because what we do not know or understand cannot be accessed in any direct or immediate way. Or again, because all of our ideas and beliefs are themselves mere images or metaphors, whose original is not immediately available.

This first clarification leads to a second one, related to the scope of Plato's conception of ἀνάμνησις. According to Socrates, the notion of recollection concerns not only the equal, the great or the less, but also beauty in itself, the good in itself, the just, the holy, 'and, in short, . . . all those things which we stamp with the seal of "being in itself"' (αὐτὸ ὃ ἔστι).¹⁹ But this does not mean that the notion of ἀνάμνησις refers only to a small part of our cognitive horizon, leaving out other, less abstract concepts. Again, these specific εἶδη are singled out because they are the basic ingredients of human beliefs, of the various combinations of human beliefs, and indeed of the whole web of assumptions underlying human cognition. Every mental representation, whatever its nature or content, entails a movement of recollection; every idea or belief 'yearns to be something else but falls short of it.'

¹⁸ In the *Phaedrus*, Socrates reproaches his interlocutor's concern for the reliability of a myth's sources: 'They used to say, my friend, that the words of the oak in the holy place of Zeus at Dodona were the first prophetic utterances. The people of that time, not being so wise as you young folks, were content in their simplicity to hear an oak or a rock, provided only it spoke the truth; but to you, perhaps, it makes a difference who the speaker is and where he comes from, for you do not consider only whether his words are true or not.' (*Phdr* 275b4–c2)

¹⁹ *Phd* 75d1–2

Even though the notion of recollection is not explicitly discussed in the allegory of the cave, the prisoners' situation is no less representative of this peculiar phenomenon. Just as the soul forgot what it knew upon entering the body, at the moment of birth, so too are the prisoners chained to their seats 'since childhood'; and just as each of the soul's ideas is a mere image or echo of a true idea, so too are the prisoners forced to see images or shadows of real things. In both cases, consciousness is burdened with a trace, or a shadow, or a memory of truth, but denied its actual content.

4.2 Towards the Light

After describing the prisoners' situation, Socrates goes on to imagine what would happen if one of them were released from his chains and made to stand up and look around.²⁰ Seeing his own body and the surrounding space, he would come into contact with a whole new reality. His cognitive horizon would no longer be reduced to the wall in front of him and the shadows would eventually be perceived as shadows and lose their referential value. However, as the narrative soon makes clear, this release would not be a full release. By getting rid of his chains, the prisoner would indeed overcome the first and most immediate form of imprisonment, but only to find himself restrained in new ways. Instead of a binary scheme, designed to illustrate a simple contrast between captivity and freedom, or ignorance and clairvoyance, Plato's allegory depicts an entire cognitive scale, made up of different and potentially endless layers of imprisonment and ignorance.

To begin with, Socrates emphasizes the idea of distance. His cave is a subterranean cave, far removed from the world up above. Its entrance is separated from the place where the prisoners live by a long uphill path, accessible only to the strongest and most persistent climbers. He points out, moreover, that this subterranean prison is very poorly lit. Its source of light is not the cave's entrance, located too far away, but a fire burning behind the prisoners, which is itself located at a great distance. Given these constraints, a prisoner who would manage to break his chains would still be faced with a number of obstacles. Having succeeded in abandoning his seat, starting up the cave's path and approaching the fire, his previous paralysis would be replaced by a new form of paralysis, caused by the fire's glare. Used to the cave's obscurity and to the shadows paraded on the wall, he would be blinded by the light and prevented from moving further. And though his new standpoint would be 'closer to reality', and 'turned more towards real things',²¹ he would be led to think otherwise. With sore eyes and a spinning head, he would 'turn and run back to what he can see, thinking that this was really clearer than what he was now being shown'.²²

²⁰See *Re* 515c4ff.

²¹*Re* 515d2–3

²²*Re* 515e2–4. Translation modified.

With time, however, the prisoner might grow used to the light and to his new surroundings. He would then manage to move further along the cave's path and approach the objects whose shadows are projected on the wall. But this new discovery would only amount to a more sophisticated re-enactment of the allegory's initial deception. Instead of shadows, the prisoner would now see 'props of all kinds', including 'statues and other creatures made of wood and stone.'²³ These objects, held by 'puppeteers' above a screen, modelled after the things that exist in the world up above, are no longer mere projections or ghostly visions, but real bodies, endowed with a specific size, colour and texture. Yet they are still only images of real things. Although the prisoner's initial ignorance would have been transformed, it would not have been eliminated. His idea of reality would still be based on a mere representation of reality, passed off as reality itself.²⁴

This second stage leads to a third allegorical situation, which takes place at a much higher altitude. If the prisoner, overcoming the first two obstacles, were forced up the cave's ascending path, leaving behind its darkened depths, he might end up reaching the cave's exit. But if he were led outside and made to look around, he would once again be paralyzed. The light of the sun, much stronger than that of the fire inside the cave, would dazzle him and prevent him from seeing anything whatsoever. Once again, he would feel dizzy and confused. The outside world would appear less real than the cave's semi-obscurity and his first impulse would be to run back inside. With time, however, he might adapt to his new dwelling and come to endure the world's brightness. But even then, he would not be immediately able to face the objects and shapes in front of him. He would start by focusing on the shadows of real objects. Afterwards, as his eyes grew less sore, he would look at the reflections of objects in water, making out their colour and shape. Finally, fully accustomed to his new situation, he would come to stare directly at things and people, discerning at last their actual appearance.

One of the most striking aspects of this description is indeed the length of the prisoner's journey and the amount of obstacles he has to face. From his initial standpoint to his final destination, he is confined to obscurity and dazzled by strong lights, led up a steep pathway and subjected to multiple forms of deception. Moreover, he is made to do these things against his will, hindered by a natural aversion to novelty and change. And yet despite the complexity of Plato's description, the allegory is not about a loose assortment of cognitive challenges, or the acquisition of different kinds of knowledge. Throughout his entire journey, the prisoner yearns for, and is repeatedly denied, *the exact same thing*, i.e. access to

²³ *Re* 514b8–515a1

²⁴ As the prisoner moves beyond the fire and encounters the screen and the objects held above it, he is unaware of the puppeteers hidden behind it. Therefore, when he hears their voices, he is persuaded that the objects themselves are speaking. As the allegory progresses, however, this illusion is not explicitly addressed. Although the prisoner eventually realizes that the world is not reduced to a puppet show, he does not meet or confront the puppeteers. Socrates' narrative focuses instead on the growing tension between light and darkness, which will culminate in the discovery of the outside world.

reality *as it actually is*. With each new trial, he realizes that the knowledge he relies on is a mere illusion, and that the things he takes to be real are in fact mere shadows, or copies, or reflections of real things. The diversity of situations imagined by Plato is not meant to illustrate the diversity of human knowledge, but the diversity of human ignorance. It is meant to emphasize the plasticity of our usual standpoint and the variety of costumes it is willing to put on to conceal its usual groundlessness.

In other words, the allegory is not so much about the alternative between ignorance and knowledge as it is about the ubiquitousness of ignorance and the possibility of actual knowledge. Whenever the prisoner believes he has overcome the realm of illusion and found reality ‘in itself’, he is eventually led to recognize that the latter is merely a more sophisticated variation of the former. This continual exposure to error is the central theme of the allegory and the main source of the imprisonment highlighted by Plato. Since our usual ignorance is characterized precisely by its appearance of knowledge, truth cannot be attained in a simple or direct way. A new account of reality, dogmatically deemed truer or more real than the previous one, is still as open to criticism as the previous one, and still as subject to being revealed as a form of ignorance in disguise. Truth can only be found, if at all, when all forms of deception have been tested and overcome. It can only emerge at the end of a long pathway of ignorance, where every intermediate stop appears at first as the finishing line. The progression outlined by Socrates is thus a *negative* progression, whose highest term is both present and absent in all of its stages. The truth supposed to lie at the end of the cave’s ascending path is already known from the very beginning, but in a negative form—as a mere assumption, or a mere signpost, or a mere yearning. The dialectical movement illustrated throughout the allegory is the movement of *δοξάζειν*, or *μετέχειν*, or *ἀναμνησκέσθαι*.

This point is crucial and often misunderstood. When reading Plato’s allegory, we tend to adopt the standpoint of a spectator and to look at the whole situation from above. Unlike the prisoners, we know that the different stops are the intermediate stages of a global progression, and we also know how it ends. Therefore, when Socrates declares that the prisoners are like us, we picture ourselves at the lowest end of a similar progression, starting with our current standpoint and leading all the way to truth. In doing so, however, we tend to conceive this new progression in a *positive* or *cumulative* way, as a succession of cognitive upgrades culminating in absolute wisdom or clairvoyance. While our usual standpoint is limited and flawed, confined to a very small portion of reality, the insight to be gained by leaving our imaginary cave is nothing less than a perfectly clear account of reality.

This kind of reading agrees with most interpretations of the so-called Theory of Ideas. Just as Plato’s *εἶδη* are usually construed as ‘enhanced’ versions of familiar concepts, so too is the allegory of the cave usually read as the illustration of an ascent towards a total or absolute cognitive perspective. As argued earlier, however, this kind of reading underestimates the seriousness of Plato’s criticism. By comparing human life to the life inside the cave, he is not claiming simply that our usual cognitive perspective is limited or flawed, that it needs expanding or perfecting. He is suggesting, moreover, that none of the things we claim to know or understand are actually known or understood by us. And he is highlighting, therefore, the need for a complete reform of human cognition.

From the prisoners' perspective, the different versions of reality adopted throughout the allegory are not the intermediate stages of a wider ascent. Each is perceived as the end of the path, i.e. reality itself. Likewise, when each version is called into question and shown to be unreal, the prisoners' disillusionment is not partial, but complete. What they realize is not that *some* of the things they believed in are false or unclear, nor that *most* of the things they believed in are false or unclear, but that *none* of them are what they seemed to be. With each transformation, the prisoners renounce their whole world and enter a new one, which they take once again to be real.

This dialectical process is illustrated by two complementary themes, namely the visual contrast between light and shadow (or between clarity and obscurity) and the physical contrast between the bottom of the cave and the world up above (or between the prisoner's ascending and descending movements). Regarding the first theme, Socrates distinguishes between 'two ways in which the eyes become confused', namely when one moves 'from darkness into light' and 'from light into darkness'.²⁵ The former movement corresponds to the prisoner's ascending journey. In the bottom of the cave, the lights are at their lowest. The fire is reduced to a distant glow and the prisoners can hardly make out the shadows on the wall. However, darkness is never absolute because the prisoners are never completely devoid of cognitive insight. Even at the lowest of cognitive stages, their outlook on reality is already a mixture of knowledge and ignorance, or vision and blindness, in the peculiar sense we have considered. Since they are already burdened with a recollection of knowledge, what they see is already a version of reality, but one that falls short of reality itself.

As one of the prisoners breaks free and moves towards the fire, the shadows recede and the light increases. Paradoxically, this change ends up compromising his ability to see the objects in front of him. Although his previous standpoint was less revealing than his new one, his vision is now blocked by the fire's glare. His first instinct is thus to return to the cave's obscurity, which he regards as clearer, and hence more real.²⁶ This situation echoes the cognitive challenge usually faced by Socrates' interlocutors throughout his enquiries: as their beliefs are examined and called into question, they are also left in a state of perplexity; as Socrates dismantles the web of assumptions that supports their usual cognitive perspective, their first instinct is to 'take a runaway's leave'²⁷ and regain their usual self-confidence.

The second source of visual confusion, experienced when one moves 'from light into darkness', is illustrated by an inversion of the allegory's initial situation. After imagining what would happen if a prisoner were to break free and exit the cave, Socrates pictures that same prisoner returning to the bottom of the cave and

²⁵ *Re* 518a2–3

²⁶ *Re* 515d–e

²⁷ *Sy* 216b5

reoccupying his initial position. This time, his eyes would be accustomed to the brightness of the sun and ‘filled with darkness’ upon entering the cave.²⁸ Having once been a foreigner in the outside world, new to the appearance and meaning of real objects, he would now be a foreigner inside the cave, incapable of making out the shadows on the wall. Before his eyes had had the time to readjust to the cave’s darkness, he would be criticized and made fun of by the other prisoners. His visual impairment would be regarded as a symptom of ignorance and foolishness, and his praise of the upper world, of its brightness and clarity, would be greeted with scepticism. Furthermore, if the liberated prisoner were to try and free his companions and convince them to follow him upwards, the latter’s initial reserve might turn into open hostility. Angered by his provocations, the cave dwellers might even decide to ‘get their hands’ on the newcomer and put him to death.²⁹

This description is also evocative of the behaviour that is usually adopted by Socrates’ interlocutors. Whereas the visual impairment experienced during the prisoner’s ascent referred to our usual cognitive standpoint, this second form of impairment is characteristic of a philosophical standpoint. In Plato’s dialogues, it is usually illustrated by Socrates’ detachment from the social and political world, his unwillingness to engage in mundane affairs and his devaluation of the things that are cherished and pursued by most people. From his point of view, human pursuits and aspirations are grounded in mere assumptions, or mere shadows, and belong therefore in the bottom of the cave. For regular men and women, on the contrary, Socrates’ concerns are themselves shady and obscure. His obsession with clarity and coherence is not only ridiculous and impractical, but also a nuisance to other people and the ultimate cause of his downfall. Like Callicles’ warnings against the practical dangers of philosophy, Socrates’ emphasis on the hostility of the cave dwellers towards the liberated prisoner is a direct allusion to his own tragic fate.³⁰

Let us now look at the second of the two aforementioned allegorical themes, namely the cave’s topography and the significance of the variations in altitude considered by Socrates. Here, it is worth highlighting two central aspects of his description. Firstly, as in the *Phaedrus*’ allegory of the flying chariot, there is a fundamental asymmetry between the ‘way up’ (ἄνω) and the ‘way down’ (κάτω), which conditions the whole allegorical setting. On the one hand, just as the soul’s winged horses were naturally destined to fly upwards, so too are the prisoners destined to inhabit the real world. They are stranded outside their natural element and prevented from fulfilling their cognitive vocation. On the other hand, just as the soul’s chariot was naturally pulled down by the force of gravity, so too are the prisoners hindered by the steepness of the cave’s ascending path. It is much easier for them to stay seated than to venture upwards, towards an unknown destination.

²⁸ *Re* 516e4–6

²⁹ *Re* 517a. Cf. *Go* 486a–c, 521b–522e, and *Ap* 31a.

³⁰ *Re* 517a4–6. See Blumenberg, *Höhlenausgänge*, 95.

In both allegories, Plato emphasizes the contrast between consciousness' 'aerial' calling and its 'earthbound' or 'sedentary' nature. Even though our usual standpoint is always aimed at knowledge, our concern for clarity and coherence is not very high. Although we are naturally interested in truth, we have a tendency to 'sit back' (θᾶκος) and take the truthfulness of what we see for granted.³¹

But the theme of altitude is also revealing of a second and equally fundamental idea. As already seen, and in line with many other Platonic texts, the allegory of the cave illustrates the contrast between ordinary cognition and the possibility of a radically different cognitive perspective. Unlike other texts, however, the allegory is particularly emphatic about the length of the journey leading up to this transformation. Given the remarkable versatility of our usual standpoint, none of the cognitive conquests made inside the cave can be taken for granted. All of them end up revealing new and more sophisticated forms of imprisonment. But not all of them are alike. The cave's path is an ascending path, starting at the bottom and ending at the top. It illustrates, therefore, a global hierarchy of ignorance, with each new conquest bringing the prisoners 'a little closer to reality'. Although each disillusionment is a complete disillusionment, each new starting point is superior to the previous one. Moreover, although the journey is very long, the prisoners are bound for a specific destination, located at a finite distance. Their fate is not to go on climbing indefinitely, as though the cave were indeed the whole world, but to reach the world outside.

Yet a word of caution is needed. As we have already seen, Plato's allegorical device is a 'Heraclitean' device: he focuses on the transition from the prisoners' standpoint to our own standpoint in order to convey the possibility of a transition from the latter to a new and even higher standpoint, located beyond our cognitive horizon. Accordingly, when the prisoners reach the end of their journey, their cognitive perspective coincides with our usual cognitive perspective, prior to the journey we ourselves are called upon to undertake. But what characterizes our usual outlook on reality is precisely the conviction that no such journey is needed—that

³¹ A similar theme is explored in the *Phaedo* (109a–114c), where Socrates imagines a group of creatures dwelling in the bottom of the sea. By means of a Heraclitean thought pattern identical to that of the *Republic*, he compares these creatures to us: like them, 'we dwell in a hollow of the earth and think we dwell on its upper surface', and 'by reason of feebleness and sluggishness, we are unable to attain the upper surface of the air; for if anyone should come to the top . . . , he could lift his head above it and see, as fishes lift their heads out of the water and see the things in our world, so he would see things in that upper world; and, if his nature were strong enough to bear the sight, he would recognise that that is the real heaven and the real light and the real earth.' (109d6–110a1) In the *Phaedrus*' allegory, human consciousness' natural 'sluggishness' was illustrated by the chariot's natural tendency to fall; in the allegory of the cave, by the steepness of the cave's ascending path and the prisoners' inability to endure the brightness of the sun. In the *Phaedo*, by comparing human beings to fish, Socrates underlines once again their natural aversion to a higher or more demanding cognitive standpoint: although his submarine beings can rise above water, they quickly lose their breath.

the world we see is the real world, made of real people and real objects. Therefore, if this conviction is unwarranted, the prisoners' final release may also be misleading. Just as the shadows on the wall, the props above the screen, the shadows outside and the reflections on water were all revealed as images devoid of content, so too can the sunlit world be part of a new and even more elaborate cognitive hoax. When the prisoners reach the end of their journey, they may be entering a second cave, lit by a new and more potent fire. And this cave may lead in turn to a third one, contained within a fourth one, and so on indefinitely. Plato's story of captivity and release may be an endless story, in which every conclusion is a new beginning and every reality the image of something else.

4.3 The Self as Shadow

Before abandoning Plato's cave, one last remark concerning shadows and real objects is in order. At the beginning of the allegory, Socrates compares our usual standpoint to that of prisoners bound by shackles and forced to see shadows projected on a wall. But although these shadows are paraded in front of us, like images on a theatre screen, the allegory does not refer merely to an external problem. The prisoners, the shadows and the cave are meant to illustrate not only what we are able to see or know about the world around us, but also *who we are* and *who we claim to be*. Plato's description is just as much a diagnosis of human cognition as it is a diagnosis of human identity and a reflection about the possibility of a radical ontological transformation.

According to Socrates, our usual account of reality is based on shadows, rather than real things. But the shadows we rely upon, unlike regular shadows, are unable to reveal the objects and events they are shadows of. Their transitivity is compromised because the things they refer to are missing, or were never known. And this problem affects not only *some* of our ideas and beliefs, nor indeed *most* of our ideas and beliefs, but *all* of them. As seen earlier, Plato's indictment is directed at the entire edifice of human cognition: from the simplest to the most complex insights, from the way we represent the surrounding world to the way we represent ourselves, and even to the way we represent our representation of the world and ourselves. Every one of our ideas amounts to a shadow, or to the shadow of a shadow, or to an even more complex combination of shadows. No matter how far or how deep we look, the actual source of our beliefs remains hidden and out of reach.

As the dialogues demonstrate, most of us tend to reject this diagnosis or to underestimate its seriousness. Like the prisoners inside the cave, we are naturally attached to our usual standpoint and naturally averse to change. Moreover, even when we do consider Plato's views and admit to being *confronted* with shadows, or *in possession* of shadow-like representations, or anything of this kind, we are led to construe the problem in a strictly objective manner. As the world's identity is called

into question, we take refuge in our own identity, as though it were immune to Plato's charges. By reducing the issue to a disagreement between selfhood and otherness, we hang on to a primary core of self-identity, unaffected by the circularity that contaminates everything else.

This expedient, for all its convenience, is still grounded in a mere assumption. Our own identity appears to be more familiar than the identity of objects and events, but it is no less bound by the rule of μέθεξις or ἀνάμνησις. Indeed, even though the idea of selfhood is undoubtedly different from the idea of a chair, or of justice, or of the greatest good, it is nonetheless mediated by the same cognitive alphabet employed to make sense of all these things. And since this alphabet, according to Plato, is entirely made up of shadows, self-identity is itself made up of shadows. Ultimately, then, by comparing our life to that of the prisoners, he is not suggesting simply that we *see* shadows, but also that we *are* shadows—and *nothing more than* shadows. On closer inspection, the problem of not knowing what we think we know is also the problem of not being who we think we are.

The *Republic's* allegory draws on an ancient literary topos, frequently found in Greek poetry and drama. In Pindar's *Pythian Odes*, for instance, man is said to be 'the dream of a shadow', set alight with 'the brilliance given by Zeus'³²; In Sophocles' *Ajax*, Odysseus declares that 'all we who live are nothing more than phantoms or fleeting shadows'.³³ These and like images anticipate Plato's own characterization of human beings as shadows of real beings, and of human existence as a shadow-like existence. In the allegory, he suggests not only that every human belief 'yearns to be something else and falls short of it', but also that every human being 'yearns to be someone else and falls short of it.' He suggests, in other words, that human life is globally burdened with the recollection of a different kind of life—a *real* life, where shadows are no longer shadows, but real objects, and beliefs no longer beliefs, but true ideas.

Once again, however, this difference should not be interpreted in positive or cumulative terms. This new life is not simply a better or enhanced version of our current life. In fact, strictly speaking, it is not new at all. The life we fall short of is the life we already claim to know and lead every day. The ontological tension Plato keeps bringing out refers not to the contrast between who we are and who we could become, in some ideal or transcendent world, but to the immanent contrast between who we are and who we take ourselves to be.³⁴ To 'know ourselves', as Socrates recommends, is thus to acknowledge that we have yet to become who we think we are and to arrive at the place where we claim to be.

³²Pindar, *Pythian Odes*, 8, l. 95

³³Sophocles, *Ajax*, l. 126. Εἰδῶλον may refer to the phantom or ghost of a living being (the εἰδῶλα wandering through Hades), but also to a mental image (see *Phd* 66c or *Th* 150c) or to an image reflected in a mirror or in water (see *So* 266b or *Re* 516a). In all these cases, the defining trait of εἰδῶλα is their *defective* nature. They amount to φαυλότερα ὄντα, i.e. 'paltry' or 'second-rate' versions of real beings.

³⁴Drew Hyland (*Finitude and Transcendence in the Platonic Dialogues*) speaks of the 'finite transcendence' evoked throughout the dialogues.

The idea that human beings are mere shadows (σκιαί), or phantoms (εἰδωλα), or images (εἰκόνες) of themselves—evoked in different contexts by Pindar, Sophocles and Plato—is also related to another important Greek topos, explicitly referenced in the *Republic*'s allegory. As Socrates highlights the incompatibility of the life led by the cave dwellers and the one that is eventually embraced by the liberated prisoner, he observes that the latter, having grown accustomed to the outside world, would lose all interest in the honours and rewards cherished and fought for inside the cave.³⁵ To illustrate this point, Socrates quotes from Homer's *Odyssey*, and more precisely from the episode where Odysseus, in search of Tiresias, descends into Hades, 'where dwell the unheeding dead, the phantoms of men outworn.'³⁶ Moving past dead servants, heroes, kings and relatives, all reduced to mere images of their former selves, Odysseus comes across the shadow of Achilles, who complains to him about his current condition. Odysseus tries to comfort him by claiming that even in death he is still a great prince, but Achilles retorts that he would rather be a poor slave to a poor master 'and live on earth' than 'to be lord over all the dead that have perished.'³⁷ These same words are used by Socrates to describe the liberated prisoner's reluctance to return to his previous life: like Achilles, he would eagerly wish to be a poor slave to a poor master and 'undergo anything rather than what he thought about down there and live in the way he did down there.'³⁸

This comparison reiterates a parallel suggested ever since the beginning of the allegory: the contrast between shadows and real things mirrors the contrast between the world of the dead and the world of the living, and the cave evokes the subterranean depths of Hades.³⁹ The implication, of course, is that the cave dwellers, reduced as they are to images or shadows of themselves, are not truly alive. Their peculiar kind of existence is grounded in the peculiar kind of transitivity we have already encountered at different junctures of the allegory: although the prisoners are not truly alive, they are not truly dead either; their life is permeated by death (or by darkness, or by deception), just as their death is permeated by life (or by light, or by truth). And something very similar happens in Homer's description: although the dead have kept their former appearance, and therefore *seem* to be alive, they were divested of their former strength, power and freedom, and reduced to mere shadows; everything in them is a sign of something else, which no longer means what it used to. They are pure *evocations*, pure *recollections* of beings whose original form is no longer present.

³⁵ See *Re* 516c-e

³⁶ *Odyssey* XI, 475–476

³⁷ *Odyssey* XI, 489–491

³⁸ *Re* 516d6–7

³⁹ See Blumenberg, *Höhlenausgänge*, 48 f.; Klär, 'Die Schatten im Höhlengleichnis und die Sophisten im Homerischen Hades'.

Like Homer's phantoms, the prisoners' power is reduced to the appearance of power, and their freedom to the appearance of freedom. They are not real people, in the true sense of the word, but mere shadows or visions thereof. However, since we readers are the allegory's true protagonists, it is our own power, our own freedom, and indeed our whole life that is being called into question: just as the cave dwellers may be dead and not know it, *we* too may be dead and not know it; and just as the cave itself may be the world of the dead, *our own world* may be the world of the dead. This possibility is explicitly mentioned in the *Gorgias*, during Socrates' debate with Callicles: quoting from one of Euripides' lost plays, Socrates wonders whether 'to live is to be dead, and to be dead, to live.'⁴⁰ This brief formula encapsulates the *Republic*'s entire allegorical situation. It aims to highlight, once again, a possibility that transcends our usual standpoint: if life as we know it is indeed a form of death, the contrast between life and death is transposed to a higher plane, in which both terms acquire a radically new significance; just as our own life is qualitatively superior to death, in the traditional sense of the word, so too a real or genuine kind of life, as Plato understands it, is qualitatively superior to the life we lead. This new life is at first merely the negative image of our current life, and can only be rendered positive through a complete reform of our usual cognitive perspective.

⁴⁰Go 492e. Socrates is quoting from Euripides' *Phrixus*, fragment 833.

Chapter 5

Philosophy and Deliverance



Abstract We have seen, so far, that the cognitive deliverance envisaged by Plato entails a complete cognitive transformation, bound to affect every aspect of human life. In this chapter I discuss the nature, relevance and possibility of this transformation, so as to determine whether Plato's scientific project is bound to remain a critical project, focused on the negative effort of uncovering reason's inner contradictions, or whether it can lead to the attainment of a truly epistemic account of reality.

I start by discussing, in Sect. 5.1, the actual need for the transformation advocated in the dialogues, as opposed to simply accepting or disregarding the doxastic nature of human cognition. Apart from Plato's discussion of this issue, I also draw on Hume's 'Platonic' approach to the subject, in order to clarify its main implications. In Sect. 5.2, I consider the extent to which the deliverance envisaged in the dialogues is actually possible. To do so, I determine and discuss the various conditions which would have to be met in order for Plato's philosophical project to be fully accomplished. In Sect. 5.3, turning to the historical legacy of Plato's thought, I argue that Hegel's idea of a dialectical re-education of natural consciousness can be read as a response to the challenge issued throughout the dialogues. Drawing on this parallel, I discuss the common ground shared by both projects and their commitment to the attainment of freedom.

Keywords Plato · Hegel · Hume · Freedom · Truth · Doxa · Science

In the previous chapters, following Plato, I highlighted the universal scope of human beings' usual cognitive imprisonment. Although personal freedom is limited in a variety of external or contingent ways, the kind of captivity described in the *Republic* and elsewhere is not external or contingent, but an essential feature of human identity. Likewise, although personal freedom may be conquered or increased in different ways, the deliverance envisaged in the dialogues is not a partial deliverance. It entails a complete cognitive transformation, bound to affect every aspect of

human life. In the following sections, I will discuss the nature, relevance and possibility of this transformation, in an attempt to determine whether Plato's philosophical project is bound to remain a critical project, or whether it can lead to the actual attainment of a true account of reality.

In either case, Plato is well aware of the obstacles faced by any serious attempt to reform human cognition. Most human beings, in order to perform the intellectual and practical tasks required of them, tend to embrace an automatic, non-philosophical standpoint, and to stand by it. For a good illustration of this, one must simply recall the reactions of Socrates' interlocutors throughout his enquiries. At first, the discovery that the most basic human concepts, when looked at more carefully, are in fact dogmatic and unclear is a natural cause of perplexity. But as the dialogues unfold, this initial perplexity gives way to exasperation, and Socrates' charges are resisted with increasing vehemence. It rapidly becomes clear that what is taking place is not a disinterested discussion about knowledge and truth, but a struggle between two conflicting modes of existence. Psychologically, by accusing his interlocutors of not knowing what they think they know, Socrates strikes a blow at their self-esteem; philosophically, his attack is much more serious, for it aims at the very core of their cognitive perspective, challenging their ability to think, choose and act the way they do.

To borrow another of Plato's images, human life, in its habitual form, is a 'life of slumber', and Socrates' role is to wake human beings from their usual 'nap'.¹ Like most sleepers, however, men and women are comfortably dreaming and do not wish to be disturbed. Moreover, even when they do consent to being woken, their alertness is usually short-lived. This incapacity to stay awake—like the prisoner's incapacity to endure the brightness of the sun, in the allegory of the cave, or the fishes' incapacity to breathe outside water, in the *Phaedo's* cosmological myth²—speaks to the paradox underlying Socrates' critical efforts: despite the seriousness and the importance of his criticism, its grip upon us is less strong than the grip of our usual standpoint, and although it is possible to live our entire life 'slumbering', the same does not seem to apply to Socrates' alternative.

In Plato's words, once we agree to join Socrates in his enquiries, 'one of two things will happen: either we shall find what we are after, or we shall be less inclined to think we know what we do not know at all; and surely even that would be a recompense not to be despised.'³ But it is precisely this latter claim that we tend to disagree with. If the contradictions underlying our usual understanding of reality were perceived or lived as an inconvenience, the acknowledgement of our ignorance would indeed be a valuable recompense. This is not the case, however, because our commitment to clarity and truth is not very serious. When faced with Plato's dilemma, we tend to choose a third and less complicated alternative: if we can spend our entire life dreaming, and if doing so is more comfortable and more practical than to try and wake up, why is philosophy really necessary?

¹ *Ap* 31a

² See Sect. 4.2, fn. 31.

³ *Th* 187c1–3

The debate concerning the need for a philosophical reform of human cognition leads to a second and even thornier one, concerning the actual possibility of such a reform: if the doxastic foundation of human beliefs is really as ubiquitous as Plato suggests, there seems to be no way of guaranteeing, beyond a reasonable doubt, that philosophy's critical charges are not themselves doxastic, or that the transformation suggested in the dialogues is not simply the more sophisticated dream of a more sophisticated slumberer. And this suspicion challenges not only the epistemic ideal pursued by philosophy, but also the freedom it is supposed to provide: if it turned out to be justified, the deliverance envisaged by Plato would amount to a new form of imprisonment, open to new forms of criticism.

5.1 The Need for Deliverance

The issues of the need for a philosophical reform of human cognition and the possibility of such a reform have haunted, in varying degrees, every post-Platonic attempt at a definition of philosophy. In modern times, David Hume addressed these issues with remarkable directness, in an effort to determine the true role played by philosophy in human life.⁴ Like Plato, Hume highlighted the basic contradiction between our cognitive self-confidence and the severe limitations of our cognitive situation. Moreover, he also pointed out that the usual success of our practical achievements is dependent on a considerable degree of ignorance and inattention, which is hardly ever problematized.

Hume's philosophical method is not that different from Plato's. In both cases, philosophy's first task is to probe the edifice of human reason, isolate its most basic components and question their meaning and intelligibility. Zooming in on elementary notions like identity, difference, cause or effect, the philosopher brings out their intrinsic contradictions and challenges their familiarity. Furthermore, insofar as these simple notions provide the basis for all other ideas and judgements, the philosopher ends up questioning the whole fabric of human reason, as well as every dimension of human life.

Hume's description of the practical and intellectual turmoil generated by this kind of criticism is reminiscent of Alcibiades' description of the effects produced by Socrates' discourses:

The *intense* view of these manifold contradictions and imperfections in human reason has so wrought upon me, and heated my brain, that I am ready to reject all belief and reasoning, and can look upon no opinion even as more probable or likely than another. Where am I, or what? From what causes do I derive my existence, and to what condition shall I return? Whose favour shall I court, and whose anger must I dread? What beings surround me? and on whom have I any influence, or who have any influence on me? I am confounded with all

⁴See Hume, *A Treatise of Human Nature*, 171ff.

these questions, and begin to fancy myself in the most deplorable condition imaginable, inviron'd with the deepest darkness, and utterly depriv'd of the use of every member and faculty.⁵

But this confusion is only temporary. After stressing the confounding power of philosophy, Hume acknowledges the lack of weight carried by philosophical propositions when compared to the assumptions of ordinary reason. Although human consciousness is capable of questioning its own methods and composition in a seemingly uncompromising manner, these 'very refin'd and metaphysical [reflections] have little or no influence upon us.'⁶ The peculiar kind of awareness promoted by philosophy, even when pursued in a serious and sustained manner, is never the rule, but always the exception.

In addition, Hume suggests that this indifference amounts to a defensive mechanism, which protects us from the disorientation bound to follow a close scrutiny of our usual cognitive limitations:

Most fortunately it happens, that since reason is incapable of dispelling these clouds, nature herself suffices to that purpose, and cures me of this philosophical melancholy and delirium, either by relaxing this bent of mind, or by some avocation, and lively impression of my senses, which obliterate all these chimeras.⁷

In response to this paradox, Hume discusses the need for philosophy by asking whether one should oppose reason's natural 'slumber', and reject 'all the trivial suggestions of the fancy'⁸, or whether it is possible to quash one's philosophical aspirations by determining, once and for all, that 'no refined or elaborate reasoning is ever to be received.'⁹ Both of these strategies turn out to be problematic. The first, if pursued in a serious manner, might lead from 'melancholy and delirium' to a complete existential paralysis: by criticizing and rejecting reason's basic cognitive alphabet, one would be left with a self-destructive scepticism, which might end up compromising one's very ability to judge and act. The second strategy, albeit more compatible with the practical demands of everyday life, might also prove untenable. If pursued in a serious manner, it might lead to a break with objectivity and to the endorsement of a complete logical relativism. By 'cutting off entirely all science and philosophy', one would 'proceed upon one singular quality of the imagination, and by a parity of reason . . . embrace all of them.'¹⁰ In the absence of a fixed criterion with which to distinguish true beliefs from false ones, reason would be reduced to self-confidence and caprice, and even the wildest absurdities could come to be accepted and advocated.

⁵Hume, *A Treatise of Human Nature*, 175. Cf. Sy 215e–216a.

⁶Ibid.

⁷Ibid. Hume's definition of philosophy as a form of *delirium* echoes Plato's definition of philosophy as a form of madness (μανία). At stake in both cases is a deviation from consciousness' normal or healthy condition.

⁸Hume, *A Treatise of Human Nature*, 174

⁹Ibid.

¹⁰Ibid.

We are left, then, with a ‘very dangerous dilemma, which-ever way we answer it.’¹¹ The choice between ignoring the contradictions unveiled by philosophy and accepting their consequences comes down, in Hume’s words, to the choice between ‘a false reason and none at all.’¹² These alternatives are very different, but both entail a debilitating form of ignorance, liable to compromise the viability of our usual mode of cognition. This danger, however, is only truly real when these attitudes are pursued for their own sake, in a radical or unrelenting manner. But this is hardly ever the case: human consciousness is always divided between objectivity and self-interest, and ready to accommodate both of these concerns. Regarding the pursuit of philosophy, Hume highlights the pragmatism of our usual standpoint, always willing to dispel the ‘clouds’ of self-criticism in favour of more efficient or rewarding alternatives. Regarding the possibility of a complete emancipation from philosophy, he highlights the intrinsically normative nature of our usual standpoint. If our concern for coherence and objectivity were entirely ignored, the very distinction between truth and falsehood would end up losing its meaning, along with all other forms of judgement and valuation. And even if this attitude were somehow tenable, it would entail a fundamental contradiction, for the decision to break free from philosophy would itself be the result of a ‘refined reasoning’, of the very kind consciousness would be set on repudiating. In this respect, Hume is also thoroughly Platonic: on his view, the expectation of truth is a defining feature of even the least demanding cognitive perspective; even the crudest or least philosophical beliefs are still in some sense ‘refined’.¹³

This basic correlation is already conveyed by Plato’s definition of philosophy. In his writings, the words φιλοσοφία and φιλοσοφεῖν stand for an attachment to ‘wisdom’ (σοφία), that is, to the truthfulness and clarity of one’s ideas and judgements. Generally speaking, to be σοφός is to be able to cognize reality as it really is, with no distortions or limitations. But although our idea of wisdom is relatively clear, our attachment to it can assume very different forms. On the one hand, as seen in the previous chapters, none of us are indifferent to the truthfulness and to the clarity of our judgements because none of us can live or act without a specific version of how things are (and were, and will be). Whether guided by the most demanding or by the most liberal cognitive standard, we are always attached to a complex system of assumptions, whose content and articulation conditions every aspect of our life. In this general sense, all of us are already, by definition, ‘followers of wisdom’ (φιλοσοφοῦντες).¹⁴ On the other hand, we are usually only philosophers *in principle*. As we have also seen, our attachment to the truthfulness of our beliefs, albeit

¹¹ Ibid.

¹² Ibid.

¹³ Unlike Plato, though, Hume admits the possibility of a *philosophical* return to natural reason, i.e. an embracement of ignorance that is no longer naïve, but grounded in a recognition of the indispensability of common beliefs and assumptions. See Livingston, *Philosophical Melancholy and Delirium*, 17–52.

¹⁴ Sy 204a8ff.

philosophical in kind, is usually hasty and opportunistic. Truth and wisdom *as such* are not usually a direct concern of ours, and only become so in the face of practical obstacles or limitations. Whenever our web of assumptions falters or is called into question, our concern with truth is rendered explicit and intensified. Faced with an unknown or unclear object, we are forced to look for its meaning, or to question our previous idea of it. But this moment of indecision is usually brief and superficial: it only lasts until a new set of beliefs is brought in, like a band-aid, to correct the problem and re-establish the system's integrity.¹⁵

Given this complex picture, Plato's dialogues are never about the simple contrast between knowledge and ignorance, or truth and error. Through his arguments, myths and allegories, he sets forth an entire cognitive scale, ranging from the lowest degree of philosophical attachment (the lowest degree of φιλοσοφεῖν) to the highest degree of philosophical attachment (the highest degree of φιλοσοφεῖν). The interval between these two limits is the realm of human life, and Socrates' habitual task is to explore its different regions, stages or altitudes.

As the dialogues usually suggest, our position within this interval is usually low and stationary. Since most of us are 'at ease' in our daily environment, focused on concrete tasks and achievements, the philosophical reassessment of our ideas feels unnecessary and inconvenient. However, as Plato repeatedly points out, the co-implication of the scale's upper and lower limits holds throughout the entire scale. Wherever we happen to be placed, our life is always burdened with a 'recollection' of knowledge. Consequently, to ask about the need for a philosophical reform of human life is not to ask about the need for something entirely new or unknown to us. Philosophy stands for an attachment to truth that is already present, albeit incipiently, in all our beliefs, choices and decisions. We are already situated—whether we want to or not—within a global scale whose summit is actual knowledge, which means we are already implicated—whether we want to or not—in a vital quest for knowledge. The real choice, then, is whether to stop halfway and settle for spurious versions of true ideas and judgements or to embark in the ascending journey depicted in the dialogues.

The first alternative is certainly the easiest. However, it may also prove the riskiest, or even the most contrary to our interests. Since our usual ignorance extends to all our beliefs, it also affects our most important choices and decisions. Therefore, by settling for a doxastic account of reality, we may in fact be placing our whole life, our whole happiness and our whole future on a doxastic foundation. It is for this reason, more than any other, that Plato insists on the vital importance of philosophy. In his eyes, the pursuit of an epistemic account of reality is not a choice among others, but the real vocation of human life. All other activities and projects presuppose a specific assessment of what is best, or most beneficial, or most important in life, but are unable to vouch for its actual truth. Grounded in mere presuppositions,

¹⁵In *Phi* 52a–b, when Socrates includes the pleasures of knowledge (μαθημάτων ἡδοναί) in the category of pure pleasures—i.e. pleasures that do not result from the fulfilment of a need or longing—he observes, rather euphemistically, that they 'are felt only by very few.'

they are inherently exposed to the possibility of error. Philosophy, on the other contrary, precisely because it rejects all presuppositions, is the only possible means of arriving at the truth about what really matters in life. And although this acknowledgement is usually not compelling enough to turn all of us into full-time philosophers, it is certainly worth considering.

5.2 The Possibility of Deliverance

Philosophy, in its strict or stronger sense, is singled out in the dialogues as the highest and most important of human activities. Only through philosophy, it is said, are men and women able to recognize the misconceptions that plague their usual outlook on reality. And only through philosophy can they hope to leave their ignorance behind and find the truth about themselves and the surrounding world. But although the former claim is demonstrated, time and again, during Socrates' enquiries, the latter is much less obvious. For it is one thing to acknowledge that one is ignorant or imprisoned and quite another to break free.

According to Plato, when human beings 'take refuge in philosophy', they 'escape from their former selves by becoming different.'¹⁶ But this transformation is at first merely a negative transformation. Philosophy's first task is not to replace one's usual beliefs with different beliefs, or one's usual understanding of reality with a new understanding of reality. Plato's method is an immanent method, whose primary aim is simply to free consciousness from its usual pretence of knowledge.

In the *Sophist*, as he addresses the issue of παιδεία and compares different educational methods, Plato refers to this negative procedure as a kind of purge (καθαρμός), designed to liberate the soul from ignorance and self-conceit.¹⁷ Whereas more traditional forms of education consist in admonishing the ignorant for their mistaken beliefs and replacing them with new ones, imparted by the teacher, the 'cathartic' method (καθαρτική) is based on a different strategy. The traditional method, he argues, is not only authoritative, but also ineffective, since 'he who thinks himself wise would never be willing to learn any of those things in which he believes he is clever.'¹⁸ Although 'traditional' pupils may be led to revise some of their views and judgements about reality, their most in-grained assumptions are bound to prevail over the teacher's corrections. To counter this problem, the 'cathartic' method sets about to 'cast out the conceit of knowledge [ἐκβολή τῆς δόξης] in another way.'¹⁹ Its proponents start by questioning a man about the things he thinks he knows or has experience of. In doing so, 'they easily discover that his opinions [τὰς δόξας] are like those of men who wander, and in their discussions they

¹⁶ *Th* 168a5–7

¹⁷ See *So* 229a–230e

¹⁸ *So* 230a6–8

¹⁹ *So* 230b1–2

collect those opinions and compare them with one another, and by the comparison they show that they contradict one another about the same things.²⁰ Through this immanent procedure, the ignorant are led to recognize the inconsistency of their beliefs. As a result, their ‘conceit of knowledge’ is abandoned and their mind is free to consider new ideas and beliefs.²¹

In the *Republic*, following the allegory of the cave, a different version of the same method is also proposed during a debate about education.²² This time, Socrates argues that teaching cannot amount simply to inserting knowledge into an empty soul, as if ‘putting sight into the eyes of the blind.’²³ This method does not work because the soul is never truly empty or blind, but full of ideas and beliefs it is strongly attached to. The teacher must focus, rather, on ‘turning the soul around’ and ‘devising a way of making the organ see, which already has vision, but is not properly orientated or looking where it should.’²⁴ The definition of this new strategy leads Socrates and Glaucon into a longer discussion about the role of education within the state and the kinds of subjects that should be taught to its leaders. Yet after examining the relative merits of arithmetic, geometry, astronomy and music, Socrates returns to his initial argument: all these subjects are mere preludes to the true and final task of education, which is to enable the soul to ‘grasp methodically in every case what each individual thing is in and of itself.’²⁵ And this final and most basic of insights can only be attained, if at all, through a negative or immanent educational method, now properly called ‘the dialectical method’.

Unlike other forms of education, designed to replace ignorance with knowledge, dialectics aims to replace one’s alleged knowledge with a higher form of ignorance, thereby clearing the way for the acquisition of real knowledge. Whereas all other τέχναι and μαθήματα—including, in varying degrees, arithmetic, geometry, astronomy and music—build their conclusions on a pre-existing substratum of beliefs, dialectics questions the entire edifice of cognition, down to its most profound sources. The dialectician isolates, criticizes, and removes one by one the different ‘hypotheses’ (ὑποθέσεις) in which one’s habitual perspective is grounded, and ‘proceeds in this way to the actual first principle in order to be securely based.’²⁶ To describe this procedure, Socrates evokes an Orphic image that is once again linked to the theme of altitude and elevation: while the ‘eye of the soul’ is usually ‘buried in a kind of barbaric filth’, dialectics ‘quietly draws and leads it upwards.’²⁷

²⁰ *So* 168b5–8

²¹ ‘They grow angry with themselves and gentle towards others, and this is the way in which they are freed from their high and obstinate opinions about themselves.’ (*So* 230b8–c2) Note the contrast between the docility of these pupils and the murderous hostility of the *Republic*’s cave dwellers.

²² See esp. *Re* 518a–519b and 521c–535a.

²³ *Re* 518c1–2

²⁴ *Re* 518d5–7

²⁵ *Re* 533b2–3

²⁶ *Re* 533c7–d1

²⁷ *Re* 533d1–3. The same image comes up in *Re* 363d7–8.

One of the key elements at play in this passage is the notion of ὑπόθεσις, whose definition, in this context, is very different from that of the modern term ‘hypothesis’. Instead of a thesis hazarded about something, admitting to a certain degree of doubt or uncertainty, the word ὑπόθεσις, just like the word δόξα, refers to a thesis that is not lived or experienced as such, but as a sure statement.²⁸ Its meaning is thus closer to that of the words ‘assumption’ or ‘presupposition’²⁹. Moreover, like Plato’s δόξαι, ὑποθέσεις are not episodic, but universal. The soul does not have a few presuppositions about a few specific objects or events: it is ‘buried’, from top to bottom, in assumptions. Here and elsewhere, Plato suggests that human cognition, prior to the cathartic intervention of philosophy, is entirely ‘hypothetical’, i.e. entirely doxastic. Finally, just like δόξαι, ὑποθέσεις appear clear and self-evident, but are in fact ambiguous and opaque.

The philosophical re-education proposed by Plato consists in recognizing one’s cognitive assumptions *as* assumptions; and what is more, as part of a ubiquitous and ultimately unintelligible web of assumptions. The aim of dialectics is thus to identify and suspend all of reason’s ὑποθέσεις, in search of an *unhypothetical* principle (τὸ ἀνυπόθετον) in which to ground a truly scientific account of reality. But this achievement is far from simple, not only because consciousness, in its usual condition, is utterly unable to anticipate what the discovery of such a principle would amount to, but also because the journey leading up to it is anything but brief and straightforward.

Let us start from the ending. Assuming such a journey were possible, and all of one’s assumptions were indeed exposed and suspended, what would one find on the other side? Different scenarios spring to mind. First, if reason’s doxastic or hypothetic web really is ubiquitous, Plato’s *unhypothetical* principle may be nothing more than an empty place. If philosophy’s role is purely negative, its ultimate truth is not a concrete or positive truth, but merely the *untruth* of human reason. If so, its ultimate reward is not pure knowledge, but pure wonder.

Second, this absolute negativity might somehow reveal a new and *positive* reality. In this case, the emptiness unveiled by the dialectician would be instantly revealed as fullness, and indeed as the absolute fullness of ἐπιστήμη. Since what ‘blocks’ the ‘eye of the soul’ is the ‘filth’ of its own assumptions, their removal would allow for a clear vision of reality, and for an instant apprehension of ‘what each individual thing is in and of itself.’

Third, as a variation of this second alternative, it might also be argued that Plato’s dialectical method is but one half—the negative half—of his philosophical project. Accordingly, once the philosophical or dialectical chapter were over, a new,

²⁸ The word ὑπόθεσις is also used by Plato in the modern sense of ‘hypothesis’, ‘supposition’, etc. (see *Eu* 11c5; *Go* 454c4; *La* 743c5, 812a4; *Me* 86e, 87a2, 87d3, 89c3; *Pa* 127d7, 128d5, 136a, 142b–c; *Phd* 92d6; *So* 244c4; *Th* 183b3), but its meaning in the *Republic*’s sixth and seventh books is closer to that of ‘presupposition’. See Carvalho, ‘Μέθοδος ε ὑπόθεσις’, esp. 39–52; Wolfsdorf, ‘The Method ἐξ ὑποθέσεως at “Meno” 86e1–87d8’, or *Trials of Reason*, 157–161.

²⁹ A translation suggested by the word’s etymology, and preserved, for example, in the German word *Voraussetzung*.

properly epistemic chapter would ensue, where cognition would no longer be reduced to *δοξάζειν*. This new world would still have to be known and explored, but reality would now be a fully intelligible place, where one's beliefs would immediately coincide with things in themselves. Following the image offered in the allegory of the cave, this would mean that the prisoners' underground ascent would be followed by a new ascent, this time in broad daylight. Instead of a subterranean path, leading to the world they already believed to inhabit, they would now have to climb a mountainous path, leading to a new and truly positive summit.

Which of these solutions, if any, is the right one? What is the real outcome of Plato's dialectical method? Instead of offering a direct answer to this question, he highlights, in different ways, the self-contradiction inherent to all of the previous scenarios. Since the aim of dialectics is a complete transformation of human cognition, its results cannot be known prior to the transformation itself. In the absence of a fully fledged philosophical standpoint, the anticipation of what it would amount to can only be entertained *ἐξ ὑποθέσεως*, that is, by means of the very assumptions consciousness is supposed to be leaving behind.³⁰

The final answer to Plato's riddle can only be revealed, therefore, at the very edge of *δοξάζειν*, when all assumptions have been successfully identified and questioned. But before such a turning point can even be considered, a more concrete problem must be grappled with. The uncertainty regarding philosophy's ultimate destination is preceded by the uncertainty regarding the actual possibility of getting there—or indeed of ever coming close. Since our usual standpoint is continually grounded in a global account of reality, the number of concepts and beliefs that come into play in each act of cognition is virtually endless. Although this multiplicity is not apparent when we consider a given object or person, for example, the very idea of an object or a person is intrinsically dependent on a whole army of silent assumptions, without which objects and persons would not be thinkable. And these assumptions relate not only to basic notions such as existence and inexistence, selfsameness and difference, truth and falsehood, number and size, etc., but also to the definition of our interests and life-goals, of what is good and evil, beneficial and harmful, worth pursuing and avoiding, etc. Consequently, a complete dialectical reversal of human cognition would require a complete survey of all these silent elements. And even if such a titanic undertaking were somehow possible, the global purge envisaged in the dialogues would also require a sure knowledge of which and how many assumptions are implicated in each of our thoughts, a clear understanding of how they change over time and a clear awareness of how they relate to one another.

³⁰ Indeed, not even Socrates seems to have reached a fully fledged philosophical standpoint: on the day of his death, with a whole life of philosophy behind him, he still claims to be 'pursuing' or 'practising' (*ἐπιτηδεύειν*) philosophy (*Phd* 64a6). Like a regular student, he reviews his doctrines and tests the soundness of his arguments, debates and refutes as though there were still much more to learn, correct and improve. His understanding of *τὸ ἀνοπότερον* is still a thoroughly negative understanding, filtered by the distorting lens of *δοξάζειν*.

To complicate matters, a second major difficulty must also be considered. As Plato recurrently points out, one of the main peculiarities of our theses about reality is their habit of concealing their thetical nature and posing as actual facts. In the allegory of the cave, as the prisoner starts his ascent, he believes he has found the knowledge he was initially denied. However, the objects above the screen, like the shadows on the wall, are eventually revealed as mere images of real things, whose actual content lies elsewhere. This disillusionment is then repeated when the prisoner steps outside the cave and looks at the shadows of real things, and again when he turns to the reflections on water. In all of these situations, what seemed real and reliable turned out to be neither. The same kind of deception applies to Plato's dialectical purge. Given the inherently doxastic nature of human cognition, there seems to be no way of ensuring that the assumptions one has singled out were indeed suspended, and not simply replaced by new ones. For even when a given belief is shown to be spurious and unintelligible, this very revelation is supported by a conviction regarding its own truthfulness and intelligibility, which is itself open to being denounced as spurious and unintelligible. And this denouncement can be redirected, again and again, at its own result, without ever reaching a final conclusion.

Hume also called attention to this problem. Echoing Plato's suspicions, he suggested that there is ultimately no solid criterion with which to distinguish between a successful reform of ordinary reason, capable of exposing and eliminating all the silent presuppositions in which it is grounded, and the covert action of yet another subset of silent presuppositions:

Can I be sure that, in leaving all established opinions, I am following truth; and by what criterion shall I distinguish her, even if fortune should at last guide me on her foot-steps? After the most accurate and exact of my reasonings, I can give no reason why I should assent to it; and feel nothing but a *strong* propensity to consider objects *strongly* in that view, under which they appear to me. . . . Without this quality, by which the mind enlivens some ideas beyond others (which seemingly is so trivial, and so little founded on reason) we could never assent to any argument.³¹

Admittedly, the effort of questioning the cogency of our ideas and beliefs is very different from our usual willingness to accept them. Nevertheless, Hume argues, the conviction with which we undertake the former task is not substantially different from the one presiding over the latter. Even though we may commit to Plato's purge, and attempt to expose the groundlessness of all our assumptions, we might be doing so as the result of a further assumption, whose content was left uncriticized. Hume is stating, in different words, what Plato had already suggested in his subterranean image: however vigorously we try to rid ourselves of the presuppositions underlying our usual cognitive perspective, there might still be a presuppositive nucleus at work within us, whose action cannot be suspended.

³¹ Hume, *A Treatise of Human Nature*, 172

Finally, adding to the two previous difficulties, a third one must also be mentioned. Apart from the sheer number and complexity of our cognitive assumptions, and from the deceptive nature of our usual standpoint, our philosophical efforts are also at odds with the ephemeral nature of their results. Even if we were able to single out and suspend all of our assumptions, such a task would immediately prove out-dated. And for two main reasons: a) because the internal composition of our usual cognitive perspective is never the same—with time and circumstances, our account of reality is continually revised and readjusted, in response to new priorities and demands; b) because even those assumptions we have successfully isolated and suspended are never permanently gone—insofar as the circumstances that led us to adopt them in the first place go on being felt (namely, the need for practical orientation, the sway of bodily needs and desires, and so on), they are bound to slip back into consciousness and regain their former influence.

Considering all of these difficulties, the possibility of a complete deliverance from error and contradiction cannot be taken for granted. Given the peculiar nature of our usual understanding of reality, philosophy can only be construed as an open process, with no clear end in sight. The philosopher's labour, as Plato understands it, is similar to the labour of Penelope: just as her web was constantly unwoven and started anew, so too philosophy's efforts tend to lose their force and must be revived and reasserted.³² The nature of philosophical conquests is such that none is ever truly guaranteed: they are not fixed or final, but part of a wider on-going movement, made of advances and retreats, victories and setbacks. As Seneca once wrote, 'when we can never prove whether we really know a thing, we must always be learning it.'³³ And that is precisely what Plato and Socrates set out to do. Rather than replacing human ignorance with some version of knowledge, they set out to denounce and resist its permanent renewal, through an equally persistent educational effort.³⁴

5.3 Plato and Hegel

All things considered, Plato's philosophy is about questions rather than answers. His writings are the very opposite of neatly constructed treatises, offering final solutions to strictly defined issues. Although their main theme is truth and knowledge, their main focus is on doubt, illusion and contradiction. Throughout the dialogues, human

³²See *Phd* 84a.

³³Seneca, *Epistles*, 191

³⁴Of course there is a different and (presumably) final solution to the philosopher's predicament. Since ignorance, according to Plato, consists primarily in a pretence of lucidity, a lucid standpoint can only truly emerge once such a pretence has been definitively abandoned. And that can only happen once we cease to have a body, and bodily needs, and the need for practical orientation, etc.—in other words, *when we cease to be alive*. 'If pure knowledge is impossible while the body is with us, one of two things must follow, either it cannot be acquired at all or only when we are dead.' (*Phd* 66e2–4)

life is described as the meeting point of two contrary and seemingly irreconcilable movements. This basic conflict is conveyed by the contrast between freedom and slavery, but also soul and body, light and darkness, dream and wakefulness or life and death. In all of these images, Plato does not simply favour one term over the other, nor does he simply claim that one is right and the other wrong. He argues that both terms are essential and usually found together. Human life *is itself* this contradiction and human beings are what they are by partaking simultaneously of both of these elements. Men and women are originally hybrid creatures, caught halfway between reason and unreason, and their time on earth—to borrow the *Republic's* haunting formulation—resembles a ‘nocturnal day’, or a ‘day shrouded in darkness’ (νοκτερινὴ ἡμέρα).³⁵

Plato does advocate a reversal of this state of affairs. Philosophy is charged precisely with ‘bringing the soul round’ from its ‘nocturnal day’ to a ‘true day’ (ἀληθινὴ ἡμέρα)³⁶, by elevating cognition from the doxastic to the epistemic realm. But if his philosophical project is directed at the attainment of a scientific standpoint, the emphasis is placed on the possibility of such an attainment rather than on its actual concretization. The dialogues are not mere guides or recipes for the acquisition of knowledge, but reflections about whether and to what extent it can be obtained. And although Socrates anticipates and exemplifies the kind of criticism required for a complete reversal of human cognition, his enquiries are only mere samples, illustrations, isolated forays into a wider and wholly unknown territory. They reveal the different ways in which our cognitive perspective falls short of actual knowledge, but they also highlight the difficulty of changing this situation.

Historically, Plato’s critique of human cognition was the starting point of a long philosophical journey. His denouncement of the intrinsic limitations of ordinary reason opened the door to a variety of critical projects, centred on different premises, aims and methodologies. Among these projects, Hegel’s idea of a dialectical re-education of ordinary consciousness can be read, in my view, as a response to the philosophical challenge issued in the dialogues.

Even though Plato and Hegel belong to very different worlds and philosophical traditions, their works are focused on the same general task, namely the elevation of human consciousness from an uneducated, or doxastic standpoint, to an educated, or scientific one. In both cases, the plea for such an elevation follows from the acknowledgement of a basic disagreement between consciousness’ usual understanding of reality and reality itself. For Plato and Hegel, our usual standpoint is grounded in a fundamental contradiction: although we are thoroughly ignorant or uneducated beings, we usually believe otherwise; and, moreover, although our ignorance limits our freedom to choose and act according to our interests, it is not usually lived or felt as a limitation. The assumption that we are more knowledgeable than we actually are entails the assumption that we are freer, more independent and more powerful than we actually are.

³⁵ Re 521c6

³⁶ Re 521c6–7

Apart from sharing the same diagnosis and taking on the same general task, Plato and Hegel also adopt a similar methodology. Since consciousness is never a *tabula rasa*, but a vast and potentially boundless repository of beliefs, and since it is usually attached to these beliefs and unwilling to let them go, the elevation to the standpoint of science cannot be achieved simply by replacing them with new beliefs. The change must come about in an immanent way, through the explicit acknowledgement of one's usual ignorance. Only when self-confidence gives way to self-criticism is consciousness free to embrace the possibility of a different outlook on reality.

In Plato's case, this feat is to be accomplished, if at all, through the dialectical method discussed in the *Republic*, the *Sophist* and elsewhere. By collecting and comparing one's usual ideas and by realizing that they contradict each other, one is led to abandon one's usual pretence of knowledge and to embrace a new, 'enlightened' ignorance, which may lead in turn to the attainment of knowledge. In Hegel's case, the elevation to science also rests on a dialectical enquiry, whereby consciousness is also led to identify, confront and overcome its own limitations. This basic methodological approach is featured, in varying forms, in most of Hegel's works, from his early essays to his later systematic writings, but its first thorough enactment is found in the *Phenomenology of Spirit*. Unlike the later chapters of Hegel's philosophical system, outlined in the *Encyclopaedia of the Philosophical Sciences*, the *Phenomenology* is a propaedeutic work, charged with an explicitly educational task. It is targeted at everyday consciousness and at the beliefs and presuppositions that shape everyday life.

Like Plato's dialectical method, Hegel's phenomenological method is designed to reveal what appears close, self-evident and familiar as distant, obscure and strange. In the preface to the *Phenomenology*, Hegel conveys this idea in thoroughly Platonic terms:

Quite generally, the familiar, just because it is *familiar* [*das Bekannte*], is not *cognitively understood* [*erkannt*]. The commonest way in which we deceive ourselves or others about understanding is by assuming something as familiar, and accepting it on that account; with all its pros and cons, such knowing never gets anywhere, and it knows not why. Subject and object, God, Nature, Understanding, sensibility, and so on, are uncritically taken for granted as familiar, established as valid, and made into fixed points for starting and stopping. While these remain unmoved, the knowing activity goes back and forth between them, thus moving only on their surface.³⁷

If cognition is to be more than a superficial habit, consciousness must turn inwards and reassess the solidity of its usual perspective. And if this effort is to bear fruit, it must cover the entire range of presuppositions that ground ordinary reason, so as to make sure that all of their effects are duly exposed and neutralized.

³⁷PS, 18 / HW 3, 35

Only ‘the scepticism that is directed against the whole range of phenomenal consciousness . . . renders the spirit for the first time competent to examine what truth is. For it brings about a state of despair about all the so-called natural ideas, thoughts and opinions . . . with which the consciousness that sets about the examination [of truth] straight away is still filled and hampered.’³⁸

However, this task is by no means an easy task. Like Plato, Hegel is well aware of consciousness’ inner complexity, of its usual dogmatism and unique aptitude for self-deception. Therefore, like Plato’s dialectical method, Hegel’s dialectical method is not limited to the simple contrast between truth and error, or reality and illusion. The *Phenomenology* sets forth a vast cognitive scale, ranging from ordinary knowledge to scientific knowledge, and documents consciousness’ laborious journey from the scale’s lower limit all the way to the top. This long ascent, divided into a series of intermediate stages, echoes the prisoner’s ascent in the allegory of the cave, equally divided into different stages.³⁹ What matters in both progressions is not the quantitative transition from a minimum to a maximum of knowledge, as though consciousness would merely accumulate new ideas and insights as it moves along. Both progressions consist in a succession of increasingly sophisticated *versions* of truth, wrongly perceived as the real thing. At each new stage, the prisoner and the phenomenological subject believe they have arrived at the end of their journey, but each new transition ends up proving them wrong: the acknowledgement that their new account of reality is merely an image or an appearance of reality gives rise to a new ‘state of despair’ and the need to go on climbing.

The most Platonic aspect of Hegel’s progression is arguably his extreme, almost obsessive insistence on the tenacity of consciousness’ presuppositive core. In the *Phenomenology*, as in Plato’s allegory, there can be no real knowledge as long as there remains within us but a single unverified assumption. Truth can only be envisaged when reason’s entire repertoire of presuppositions (or shadows, or images, or ὑποθέσεις) has been successfully neutralized, and the *Phenomenology*’s lengthy progression is the detailed history of this struggle, extended to every corner of human cognition. But here, where Plato and Hegel come closest to each other, is also where they part company. Although the phenomenological progression is grounded in an originally Platonic idea, Hegel’s encyclopaedic fervour contrasts with Socrates’ tentative enquiries. Whilst Plato’s philosophical project is above all

³⁸PS, 50 / HW 3, 73. Regarding the translation of *erscheinendes Bewußtsein* as ‘phenomenal consciousness’, see Sect. 1.3, fn. 44. As for the ‘state of despair’ prompted by scepticism, the translation is unable to capture Hegel’s wordplay: *Verzweiflung* means despair, but it is also derived from *Zweifel*, ‘doubt’. The term also evokes, therefore, the movement of being led into a ‘state of doubt’.

³⁹See Philonenko, *Leçons Platoniciennes*, 331: ‘Platon suppose qu’“on” libère un prisonnier; ce dernier voudra s’élever à la lumière, mais plus il se rapprochera de la lumière, plus ses yeux habitués aux ombres le feront souffrir. En un sens, il est question d’une véritable *Phénoménologie de l’esprit*, car enfin ce “on” qui déchaîne l’âme doit lui-même être libre, et sa position par rapport au feu est juste l’inverse de celle du prisonnier.’ The structure of Hegel’s *Phenomenology* and the meaning of its different standpoints—viz. the standpoint of the ‘narrator’ (‘on’, or ‘we’) versus the standpoint of natural consciousness (‘es’, or ‘it’)—will be discussed in Section 6.3.

an *interrogative* and *inconclusive* project, continually haunted by the shadows of failure and unfulfilment, Hegel's phenomenological project is meant to be *affirmative* and *conclusive*, directed towards a concrete and final destination.

The development of 'philosophic science as science', Hegel writes, 'begins with Plato.'⁴⁰ But although the latter's standpoint is 'clearly defined and necessary, . . . it is impossible for us to remain there', for 'Reason now makes higher demands.'⁴¹ The entire history of philosophy can be interpreted, in Hegelian terms, as the long and fruitful development of Plato's critical diagnosis, culminating in Kant's 'return to the standpoint of Socrates', albeit 'on a higher level.'⁴² Departing from this long negative lineage, Hegel takes on the positive task of solving Plato's riddle and freeing consciousness from the fetters of ignorance and self-deceit. His work is infused with a unique philosophical hubris, whose closest match is arguably Nietzsche's own (anti-)Platonic attempt to reform human cognition. But whereas Nietzsche takes issue with the starting point of Plato's philosophy, and with dialectics in general, Hegel embraces the dialectical tradition and criticizes its ultimate lack of direction. For him, the release from consciousness' usual cognitive bondage can only come about once dialectics is raised to a systematic form. In other words, if Plato's journey is to be successful, the different stops along the way cannot be regarded merely as successive or contingently related events. They must emerge as the interconnected stages of a global progression, whose goal is as necessarily fixed for consciousness as the path leading up to it.

⁴⁰LHP, vol. 2, 1 / HW 19, 11

⁴¹LHP, vol. 2, 10 / HW 19, 20

⁴²LHP, vol. 3, 424 / HW 20, 329f.

Part II

Hegel's Phenomenological Project

Chapter 6

Hegel's Phenomenology of Spirit



Abstract Few authors have taken Plato's insistence on the need for a philosophical liberation of humankind as seriously or as far as Hegel. Although this liberating effort pervades Hegel's writings, I focus mainly on the *Phenomenology of Spirit* and its unique emphasis on the relationship between freedom, knowledge and education. Like Plato's philosophical paideia, and unlike the Hegelian system's subsequent stages, the *Phenomenology* dispenses with previous philosophical assumptions. Its starting point is human consciousness in its native or pre-philosophical form.

Focusing on the *Phenomenology*'s aims, methodology and structure, I start by suggesting, in Sect. 6.1, that Hegel's work can be read both as a development of, and a departure from, Plato's critical standpoint. Especially important in this context is the formal parallel between the ascent described in the allegory of the cave and the 'way of despair' set out in the *Phenomenology*. Afterwards, Sect. 6.2 discusses the organic nature of Hegel's progression, its thematic scope and its alleged completeness. Finally, Sect. 6.3 addresses some of the *Phenomenology*'s most controversial issues, namely the relationship between the different points of view present within the text and the exact meaning of the progression's final stage.

Keywords Hegel · Plato · *Phenomenology of Spirit* · Freedom · Truth · Dialectic · Negation

*"Wie komm ich am besten den Berg hinan?"
Steig nur hinauf und denk nicht dran!"¹*

In line with Plato and the Platonic tradition, Hegel's conception of freedom is part of a wider philosophical survey concerning the nature, scope and value of human cognition. To be free, one must be in possession of a clear account of reality, in which one's judgements, choices and decisions can be safely grounded. According

¹GS, 14 / KSA 3, 356: 'How do I best get to the top of this hill?' / 'Climb it, don't think it, and maybe you will.'

to Hegel, however, this is precisely what human beings usually lack. Human consciousness, in its habitual form, is ignorant and self-deluded, guided by a series of assumptions whose truthfulness it takes for granted but is unable to guarantee. Consequently, human life, in its habitual form, is not a free life. Freedom can only be achieved through a global re-education of ordinary consciousness, aimed at the attainment of a true or scientific cognitive perspective.

For Hegel, as Alan Patten rightly observed, 'the problem of how an agent develops and reinforces the capacity for free and rational agency is primarily a problem of how he can arrive at and sustain the right sort of self-understanding. When an agent is unfree, or a "slave", as Hegel often puts it, this is usually due not to his chains but to his underdeveloped self-conception or self-understanding.'² Freedom requires, therefore, that one focuses not only on the external circumstances of one's decisions and actions, but also—and primarily—on an inner core of personal knowledge. It requires a reassessment of the basic structure of human cognition—of what it means to be a cognitive subject, faced with an objective world, open to different forms of interpretation.

This is the task of the *Phenomenology of Spirit*. Hegel's first major opus is a highly ambitious work, with a unique structure and a wide thematic scope. It consists in a long succession of cognitive stages, leading from the most immediate mode of cognition, which Hegel names 'sense-certainty', all the way to the standpoint of science, or 'absolute knowing'. Throughout the progression, readers are led to consider different forms of self-understanding, corresponding to different outlooks on reality. What is more, they are led to examine these different standpoints, to assess their solidity and to acknowledge their inner contradictions. At first, each of them appears to offer a clear and accurate representation of human cognition, coupled with a clear and accurate representation of the objective world. However, as the progression unfolds, this appearance is invariably shown to be illusory. Each standpoint is then replaced by a new one, and the same process is repeated again and again, leading to new and increasingly sophisticated modes of cognition. Hegel sets out a complete educational programme, covering a vast array of concepts, subjects and attitudes. The complexity of his survey is meant to reflect the complexity of human life and the various cognitive directions available to human consciousness.

Yet despite this diversity, the *Phenomenology* follows a very specific plan. All the stages are connected and their meaning is determined by their position within the progression. Therefore, unlike many other philosophical texts, the *Phenomenology* cannot be broken down into a collection of arguments about a specific subject or set of subjects. Although its different sections are sometimes considered separately, in connection with concrete philosophical debates, their meaning can only be truly grasped in its original context. By the same token, the merits of any of the *Phenomenology*'s particular theses can only be fully appreciated in light of the work's global structure. Since each stage is but an intermediate stop along a wider argumentative path, its specific conclusions are necessarily provisional. They are

²Patten, *Hegel's Idea of Freedom*, 124

neither true conclusions nor completely false ones: as the progression moves forward, each stage is examined, criticized and abandoned; however, it is also revealed as the necessary vehicle for the attainment of new stages and conclusions. Ultimately, then, its truth must not be looked for in itself, but in the progression's final destination. It hinges on the cogency of Hegel's methodological approach.

In what follows I offer a brief discussion of Hegel's progression. The analysis is divided into three main parts: in the present chapter, I deal with the *Phenomenology* as a whole and discuss its aims and methodology; in Chaps. 7–9 I delve into the actual progression and examine some of its central moments; in the last sections of Chap. 9 I discuss the *Phenomenology*'s relationship to Hegel's later works and to future forms of philosophical criticism. My overall goal is not to offer a detailed reading of the entire phenomenological progression, but to focus on the issues of freedom, power and knowledge. I will highlight the different ways in which Hegel's text echoes Plato's critique of human freedom, the new elements it adds to the previous debate and the kind of liberation it proposes.

6.1 Truth and Appearance

The *Phenomenology of Spirit* is deeply indebted to the Platonic tradition. Like Plato, Hegel challenges the legitimacy of ordinary reason and highlights the need for a global reform of human cognition. Furthermore, he points out that this problem is not usually seen or experienced as such. Although we admit to being ignorant or mistaken about an endless variety of issues, we consider ourselves competent enough to know most daily objects *as they really are*.

The unwillingness to question our beliefs in a serious way is tied, once again, to the universal scope of our cognitive standpoint. Following Plato, Hegel highlights the fact that human cognition is never only about a specific region of reality, but always about the whole of it. Our knowledge is not confined to a series of islands interspersed by gaps of complete ignorance. We are guided, at any given moment, by a global version of things, presupposed and reiterated by every new act of cognition. And this global picture is not rigid, but plastic and versatile. It can be expanded, revised and adjusted to new situations. Therefore, when faced with new objects and events, we are never lost in pure wonder. They are always perceived as extensions of a familiar cognitive landscape and adapted to a familiar set of rules. Their novelty is perceived as a variation of the same cognitive repertoire, whose global legitimacy is taken for granted.

To return to a key Platonic image, our usual standpoint has its own alphabet. Its apparent virtuosity derives from the ability to combine different letters in a seemingly infinite number of ways. Yet all of the words and sentences we are able to form derive from a finite group of letters. And these letters, when subjected to a closer inspection, are not as clear as we tend to assume. For Hegel, as for Plato, a true reform of human cognition depends on the acknowledgement of this double paradox. Even though our usual perspective seems remarkably free and malleable, ready to cope with all kinds of situations, it is really self-centred and remarkably unclear. Once this contradiction is recognized, the universal enclosure of ordinary reason can

finally be breached: if the world we are used to is only a specific *version of reality*, it can be replaced by other versions of reality. Indeed, it might just be that our current cognitive alphabet is not the alphabet in which reality is actually written.

These preliminary remarks are sufficient to dispel a few misconceptions commonly associated with the *Phenomenology*. Contrary to what is sometimes assumed, Hegel's progression is not a cumulative progression. The elevation of ordinary consciousness to the stage of 'absolute knowing' does not amount to a simple increase in the scope of one's cognitive outlook, as though the phenomenological subject would start by knowing very little about reality and would end up knowing very much. The progression is not about completing an incomplete cognitive picture, or adding cognitive layers to the same cognitive core. It documents a different and more profound transformation, or series of transformations, extending to the very notions of reality, identity and cognition.

Likewise, the elevation to the stage of 'absolute knowing' does not amount to a simple increase in the clarity of one's cognitive outlook, as though the phenomenological subject would start by seeing things in a vague or unclear way and would end up seeing them very clearly. Indeed, Hegel's progression is not about sharpening an obscure version of reality, or adding detail to the same cognitive image. The transition from each stage to the next is fuelled by a different kind of transformation, with wider and more profound implications.

Both of these readings follow the aforementioned tendency, characteristic of ordinary reason, to construe difference and change in an immanent way. Using our usual standpoint as reference, we imagine a series of enhanced or amplified forms of knowledge, leading to a universal and absolutely transparent vision of reality. Yet this vision, however wide and illuminating, is still but a variation of our usual cognitive alphabet. The 'absolute knowing' it is thought to convey is merely the absolute version of a specific kind of knowing, which is still as regional and dogmatic as before, and still affected by the structural limitations Hegel aims to overcome.

Marcel Proust once wrote that 'the only true voyage . . . would be not to visit new landscapes, but to possess other eyes'.³ This is the kind of voyage set out in the *Phenomenology*. With each new stage, what changes is not merely the scope or the clarity of what is seen, but the very act of seeing. Consciousness comes to see different things, but it comes to see them differently, with the aid of new eyes and a new cognitive alphabet. Instead of exploring and expanding its initial world, it is led into a new world, or a succession of new worlds, where nothing is what it used to be. And precisely because these new worlds are not similar to the initial one, they cannot be envisaged using the latter's cognitive resources. As long as consciousness holds on to its usual eyesight, it is incapable of actual change. It can only anticipate the kind of leap advocated by Hegel in a negative way, as a blind man anticipates the world's visual appearance.

But this is still only part of the story. Although Hegel insists on the radical difference between each phenomenological perspective and the next, he also claims that all stages are nonetheless connected—that they are all somehow linked to a

³Proust, *À la Recherche du Temps Perdu*, vol. 3, 258

single overarching principle, which lives on in every one of them and holds the key to their mutual interaction. The *Phenomenology* is not a mere collection of world-views, but an orderly and organic progression, brought about ‘through one and the same principle having spontaneously assumed different shapes.’⁴

Following this claim, Hegel also argues that each different stage, because it is the product of the same generating principle, is already somehow revealing of all the others. It already contains within itself, inexplicitly, the entire progression, as well as its final result. And this also applies to the *Phenomenology*’s starting point: in Hegel’s words, truth is ‘the circle that presupposes its end as its goal, having its end also as its beginning.’⁵ Although consciousness, in its most basic form, is confined to a specific cognitive alphabet, it is nonetheless bound by a higher cognitive calling; although it is unable, from its limited vantage point, to envisage other, more sophisticated modes of cognition, it is already burdened, *in posse*, with all its future insights.

But how can each phenomenological perspective be *at the same time* a whole and a part, self-enclosed and dependent on something else? At first glance, these two claims appear irreconcilable. And yet if either of one them is favoured over the other, the whole progression, as Hegel understands it, will be compromised. On the one hand, if each standpoint is truly self-enclosed, consciousness is liable to remain eternally tied to the same cognitive alphabet. If every form of otherness is but a variation of the same cognitive repertoire, our native standpoint is indeed inescapable and no real progression can ever take place. On the other hand, if each stage is not a whole world, but a mere region within a wider cognitive standpoint, we are led back to the cumulative model that was just criticized. In this case, consciousness may indeed discover new landscapes, but it is bound to do so with the same pair of eyes. Just like before, difference and otherness are mere variations of the same familiar alphabet and no real progression can ever take place.

In the introduction, Hegel defines the entire progression as a *Realisierung des Begriffs*,⁶ that is, as the complete ‘realization’ of a single unifying ‘notion’. As he proceeds to make clear, consciousness is always burdened with an idea or ‘notion’ of reality—or rather, it *is itself* this idea or notion. Up until now I have used the terms ‘standpoint’, ‘perspective’, or ‘outlook’ interchangeably, in order to characterize consciousness’ usual *version of reality*, but there is an important difference between these expressions and Hegel’s specific definition of *Begriff*. Throughout the *Phenomenology*, he argues, consciousness is led to endorse different versions of reality—but not, strictly speaking, different ‘notions’ thereof. In the beginning, in the end and in every intermediate stage, consciousness is guided by *the same basic notion*. Its understanding of it, however, is not always the same: it depends on its *degree of realization*, which is altered with each new transition.⁷

⁴PS, 8 / HW 3, 21

⁵PS, 10 / HW 3, 23

⁶PS, 49 / HW 3, 72

⁷Throughout the introduction, Hegel’s use of the term *Begriff* is twofold. In a more general sense, which is the one that I will favour in the following pages, *Begriff* is the guiding principle of the

At the beginning of the progression, this overarching notion is fully 'unrealized'. One's ideas and judgements are thought to capture things as they really are, and one's version of reality is thus not lived or experienced as such, but as reality itself. Yet this initial agreement is only apparent: once consciousness is brought to the stand and confronted with its own judgements, it '[shows] itself to be only the *notion* of knowledge, or in other words, not to be real knowledge.'⁸ Contrary to what was assumed, our initial idea of reality does not coincide with reality itself, nor is it able to convey things as they really are. But this realization is at first merely negative. Since our cognitive perspective is a global perspective, which determines all our thoughts and judgements, it extends both to our idea of reality and to reality itself. Therefore, we cannot replace our initial version of reality with a truer one without falling back on the same cognitive alphabet and repeating the very mistake we aim to avoid.

In other words, we are trapped within a standpoint we know to be untrue and unable to escape its contradictions. To solve this problem, Plato devised a negative form of education, grounded in an immanent critique of human cognition. He argued that actual knowledge can only be attained through a complete refutation of what we already claim to know. For Hegel, this negative approach is the starting point of a long critical tradition, which links Plato to Stoicism and Scepticism, and whose basic premise is still at work in Kant's critical philosophy. In the *Critique of Pure Reason*, Kant also focuses on the disagreement between consciousness' usual account of reality and reality itself. Moreover, he also endorses an immanent form of criticism, equally motivated by the universal (or transcendental) scope of natural reason: since ordinary consciousness is confined to a global account of reality, it cannot step outside of it and assess the degree to which it coincides with reality itself. Things *in themselves* (viz. things as they are independently of one's knowledge of them) can only be cognized indirectly, as that which remains after the acknowledgement and suspension of reason's global lens. This negative procedure is made possible through a detailed exposition of consciousness' internal composition, including all of the 'secrets judgments of common reason'.⁹ Once this exposition is complete, the critic can subtract it to his or her usual account of reality and isolate reality's true nature.

This whole strategy can be translated into a simple formula:

$$R + C = R'$$

where R stands for *reality in itself*, C for the *ability to cognize reality* and R' for the joint result of these two variables, i.e. *cognized reality*.

entire progression, and that which every stage seeks (and fails) to capture. Accordingly, the *Phenomenology* as a whole amounts to a *Realisierung des Begriffs*. But Hegel also uses the term *Begriff* to designate the specific notions adopted by consciousness at each phenomenological stage. In this second sense, there is not only one *Begriff*, but different *Begriffe*, corresponding to different objects and cognitive criteria. These two uses are nonetheless closely connected: the different notions adopted by consciousness amount to different *versions* of the same founding notion. They are pseudo-notions, as it were, whose truth can only be found at the end of the progression.

⁸PS, 49 / HW 3, 72. Emphasis added.

⁹Kant's *gesammelte Schriften*, vol. 15, 180

Kant's critical project consists in an inversion of the way we tend to conceive the relationship between knowledge and reality: rather than assuming, as we normally do, that reality exists in its own right and cognition is somehow added to it, Kant points out that all reality is by definition *cognized reality*, and hence that the very idea of an object prior to any act of cognition is itself the correlate of an act of cognition. In other words, he highlights the transcendental scope of natural reason and the impossibility of attaining truth in any immediate or direct way. Instead, he proposes to reverse the operation set out in the first equation. Taking cognized reality (R') as its starting point, the critical project seeks to identify and subtract the distorting effects of human cognition (C):

$$R' - C = R$$

According to Hegel, although Kant correctly emphasized the universality of natural consciousness and the need for an immanent critique of human cognition, his critical strategy is not critical enough. Despite his so-called Copernican revolution, Kant still regards cognition and reality as two separate things, opposed to one another. His entire critique is grounded in a dichotomic approach to human cognition that presupposes the very difference it strives to overcome.

This point is made in the opening paragraphs of the introduction, where Hegel contrasts this critical approach with his own critical approach. It should be noted, however, that this initial charge is not aimed solely at Kant. Hegel blames critical philosophy for having failed to criticize a 'natural assumption' that is already at work in human beings' usual cognitive perspective. Namely, both the critical standpoint and our ordinary standpoint tend to view reason uncritically as either an instrument (*Werkzeug*) for 'getting hold of the truth' or the medium (*Mittel*) whereby the 'light of truth' reaches us. Whereas natural consciousness regards this medium as a transparent filter, the critical project acknowledges its distorting power. But although the critical model is undoubtedly more sophisticated than the natural one, both presuppose a fundamental divorce between reason and reality. In both cases, the true is defined as that which lies *beyond* cognition, and which the latter is supposed to reveal.

For Hegel, this approach is inherently contradictory. In the first case, if cognition is indeed an *instrument* used on reality, its use can either bring reality closer to us without altering its original nature—this is consciousness' ordinary, *uncritical* view—or reveal the world at the price of altering its shape—this is the philosopher's *critical* view. We may try to overcome this dilemma, as the philosopher does, by getting acquainted with the way in which this instrument works, but this effort is also doomed to fail: by removing 'from a reshaped thing what the instrument has done to it, the thing . . . becomes for us exactly what it was before.'¹⁰

¹⁰PS, 46 f. / HW 3, 69

In the second case, if cognition is the *medium* whereby reality is brought into view, it can either offer a transparent image of reality—this is the *uncritical* view—or reveal the world at the price of distorting it—this is the *critical* view. Again, we may try to overcome this dilemma by determining how this medium filters the ‘light of truth’, but this effort is equally vain. ‘If by testing cognition, which we conceive of as a *medium*, we get to know the law of its refraction, it is again useless to subtract this from the end result. For it is not the refraction of the ray, but the ray itself whereby truth reaches us, that is cognition; and if this were removed, all that would be indicated would be a pure direction, or a blank space.’¹¹ When the instrument or medium employed by human reason is removed, *everything else is removed*. Cognition and reality cannot be detached from one another because they are inherently reciprocal: every act of cognition is by definition directed at a given reality, and every reality is by definition the correlate of an act of cognition. Each is what it is *through the other*, which means that a serious quest for knowledge must consider both *simultaneously*.

The *Phenomenology*’s starting point is therefore the dialectical correlation between consciousness and its most immediate object: *das erscheinende Wissen*, i.e. *phenomenal* knowledge, or the knowledge that first *appears* to consciousness. Instead of deploying a specific kind of critique, or aiming at a specific standard of truth, Hegel takes up this correlation in its own terms. No method or aim is needed because consciousness is already ‘the notion of itself’¹²: it contains within itself the key to its own transformation, and must simply be led to recognize as much. Whereas natural or unconscious beings are confined to a fixed mode of identity—that is, whereas they *are what they are*, and nothing more—consciousness is the source of its own identity—it is what it *determines* itself to be, and can rise therefore to new and more sophisticated modes of identity.

To understand what this means, let us look closer at Hegel’s basic characterization of cognition. Whenever consciousness cognizes a given reality, it ‘simultaneously *distinguishes* itself from something and at the same time *relates* itself to it.’¹³ More precisely, it distinguishes reality *in itself*, or *truth*, from reality *for consciousness*, or its *version* of truth. Now the assessment of the truthfulness of one’s knowledge usually consists in a comparison between one’s version of reality and reality itself. However, since both of these elements fall within consciousness, this assessment amounts to a comparison of consciousness with itself. Both what is examined and the standard of examination are *for consciousness*, which means that the exam is directed simultaneously at the truthfulness of one’s thoughts and at the idea of truth itself.

Consciousness knows *something*; this object is the essence or the *in-itself*; but it is also for consciousness the *in-itself*. This is where the ambiguity of this truth enters. We see that consciousness now has two objects: one is the first *in-itself*, the second is the *being-for-*

¹¹ PS, 47 / HW 3, 69

¹² PS, 51 / HW 3, 74

¹³ PS, 52 / HW 3, 76

consciousness of this in-itself. The latter appears at first sight to be merely the reflection of consciousness into itself, i.e. what consciousness has in mind is not an object, but only its knowledge of that first object. But . . . the first object, in being known, is altered for consciousness; it ceases to be the in-itself, and becomes something that is the in-itself only for consciousness.¹⁴

In other words, consciousness is simultaneously the consciousness of its object (the ‘in-itself’) and the consciousness of its knowledge of the object (the ‘being-for-consciousness of the in-itself’). If these moments disagree with each other—if the object *in itself* is merely *for consciousness*, a subjective version of reality instead of the real thing—one’s usual standpoint is shown to be ineffective and in need of transformation.

At first, this realization is merely *negative*. If one’s version of reality does not coincide with reality itself, the problem seems to lie in the former rather than the latter. ‘It seems that consciousness cannot, as it were, get behind the object as it exists *for consciousness* so as to examine what the object is *in itself*.’¹⁵ As a result, consciousness is faced with the nothingness of its usual knowledge, and shut out from reality and truth. But this negative attitude fails to appreciate the dialectical nature of this transformation. Since consciousness is simultaneously consciousness of reality and consciousness of itself, the acknowledgement of its ignorance entails a transformation of the very criterion whereby knowledge and ignorance are measured. If ‘consciousness thus finds that its knowledge does not correspond to its object, the object itself does no stand the test’¹⁶: what seemed to be *in itself* is now only *for consciousness*, and this discovery is *itself the source of consciousness’ new object*.

Hegel refers to this movement as the *experience* of consciousness and highlights its positive significance.¹⁷ The critical tradition, by focusing exclusively on the nothingness of ordinary reason, prescribed a purely negative kind of deliverance. It failed to see that the untruth of ordinary consciousness is not just any untruth—not an absolute, generic or *abstract* untruth, but a *determinate* untruth, viz. the untruth that emerges from the negation of a determinate content. In the *Phenomenology*, on the other hand, the negation of natural knowledge is a *determinate* negation, whose positive content gives access to a higher cognitive stage. Once consciousness realizes that its knowledge is only the appearance of knowledge, the latter is itself posited as the new object. ‘This new object is the nothingness of the first, it is what experience has made of it.’¹⁸ But since consciousness is simultaneously objective and subjective, knowledge *in itself* and knowledge *for itself*, this new object also gives rise to a new criterion of truth and to a new dialectical agreement. Truth is to be measured through a new self-comparison, which leads in turn to a new disagreement and a new dialectical reversal.

¹⁴PS, 55 / HW 3, 78f.

¹⁵PS, 54 / HW 3, 78

¹⁶Ibid.

¹⁷The term ‘experience’, *Erfahrung*, is built on the verb *fahren*, i.e. to ‘go’, ‘ride’, ‘travel’, etc. Experience is thus an inherently *dynamic* process, leading to increasingly higher forms of cognition.

¹⁸PS, 55 / HW 3, 79

Whenever this cycle is renewed, a new 'shape' of consciousness comes on the scene. A new world is revealed to consciousness, along with a new cognitive alphabet and a new set of aims and priorities. But although this world is irreducible to consciousness' previous cognitive situation, it is nonetheless the necessary result of what went before, a new solution to the same founding predicament. This dialectical mixture of novelty and repetition is the key to understanding Hegel's answer to the question raised at the beginning of this section: namely, how can each phenomenological stage, be *at the same time* a whole and a part? This paradox ceases to be one when the notions of identity and difference cease to be construed in strictly dichotomic terms. Identity is never simple or immediate, but always the result of the agreement between a given version of reality and a given standard of truth. Likewise, difference is never simple or absolute, but always the inversion of this agreement, leading to the emergence of a new version of reality and a new standard of truth.

Each phenomenological transition is both a revelation and a repetition: on the one hand, each new stage introduces a different object and a different criterion of truth; on the other hand, each amounts to a more sophisticated translation of the same original notion. But although this guiding notion is present throughout the entire progression, it is not preserved in a simple or direct manner: what is maintained from stage to stage is not a specific object or a specific criterion of truth, but the very drive to overcome consciousness' internal contradictions. The negation of a given standpoint is never final because it preserves the need to accommodate its initial demands in some other way. In Hegelian language, this positive negation is called *Aufhebung*, or sublation: it is the movement whereby a given standpoint is simultaneously negated, preserved and raised to a higher form.¹⁹ Or rather, it is the movement whereby a specific worldview is negated while its underlying notion is preserved and raised to a higher degree of realization.

The *Phenomenology's* first chapter, devoted to 'sense-certainty', starts out with a simple equation: knowledge is the immediate correspondence between 'what simply is' and one's direct apprehension of it. Yet as consciousness tests this initial definition, by comparing its object with itself, a series of contradictions start to emerge. It turns out that the knowledge consciousness is so certain of is not able to account for reality, which means that the world revealed by sense-certainty is not real in itself, but only for consciousness. The sensuous immediacy ascribed to reality is not objectively true, but subjectively asserted, and the negation of this immediacy, i.e. universality, becomes consciousness' new object.

¹⁹ As even the most basic introduction to Hegel is sure to point out, the word *Aufheben* conveys both a negative meaning, which can be translated by the English verbs 'to suppress', or 'to subsume', and the positive idea of 'maintaining' or 'preserving' something. Furthermore, the radical *heben* ('to lift', 'to elevate') reinforces the idea of a transition from a hierarchically inferior to a hierarchically superior stage. However, *Aufheben* may also have other meanings. It does not always refer, in Hegel's writings, to the technical procedure usually translated as 'sublation'.

This first *Aufhebung* leads to the emergence of a new phenomenological stage, namely ‘perception’ (*Wahrnehmung*), and a new version of reality, i.e. the world of ‘things’. Once again, however, by comparing its object with itself, consciousness comes across a series of contradictions. Perception proves unable to account for the reality it purports to reveal, which means that the world of things is not real in itself, but only for consciousness. As Hegel puts it, the ‘conditioned’ universality ascribed to reality is turned into its opposite, i.e. ‘unconditioned’ universality, and posited as consciousness’ new object.

A new self-comparison and a new *Aufhebung* will lead, afterwards, to the stage of ‘understanding’ (*Verstand*), which will then pave the way for the stage of ‘self-consciousness’, and then ‘reason’, and then ‘spirit’, and finally ‘absolute knowing’. Throughout this long journey, consciousness is faced with new objects and with new criteria of truth, but all of them prove self-contradictory. According to the progression’s immanent logic, the experience of consciousness is bound to continue until its knowledge and its object no longer disagree with one another. When reality in itself is no longer merely for consciousness—that is, when one’s version of truth coincides with one’s criterion of truth—subject and object will no longer contradict each other. In this final stage, ‘knowledge no longer needs to go beyond itself’, for the ‘notion corresponds to the object and the object to the notion.’²⁰

6.2 Experience as Science

Hegel’s phenomenological critique is a double critique, directed at two major cognitive attitudes. Its first and most immediate target is ordinary reason, or the kind of knowledge held by most of us in the course of our lives. The *Phenomenology* aims to denounce and overcome the naïve dogmatism of our usual standpoint and to replace it with a true or scientific one. But Hegel’s critique is also targeted at philosophy and philosophical reason. It also aims to denounce and overcome the *cultivated* dogmatism of critical philosophy and to replace it with a truly epistemic outlook on reality.

In the first case, Hegel’s diagnosis coincides with the critical diagnosis. Usually, whenever we consider our own ignorance, we tend to construe error and untruth as regional problems, due to a simple lack of correspondence: if our understanding of reality fails to capture a given object or event, we replace it with a more complete or accurate version of things. The world remains fixed, so to speak, and it is up to consciousness to find its true image. Returning to Plato’s analogy, this model rests on the assumption that the cognitive alphabet employed by consciousness is the same in which the world is written, and that error and untruth amount to wrong words, or wrong sentences, whose letters must be combined differently. To counter this assumption, the critical tradition shifts the focus from the words and sentences to

²⁰PS, 51 / HW 3, 74

the letters themselves and questions their ultimate meaning. If the ideas and concepts that support natural reason are defined by other ideas and concepts, and if the latter's meaning is bound, in turn, to other ideas and concepts, it might just be that reason's alphabet is entirely self-referential. In Hegel's language, it might be that all our ideas are merely *for consciousness*, and not *in itself*, and hence that natural cognition is constitutively unable to reveal reality as it actually is.

For Hegel, this discovery is the necessary starting point of any philosophical enquiry. Only a scepticism 'that is directed against the whole range of phenomenal consciousness', can render 'the Spirit for the first time competent to examine what truth is.'²¹ But this discovery is not sufficient. If criticism stops here, philosophy is reduced to an insoluble contradiction: on one side, although reason is purely *for itself*, it is bound to posit a real world beyond itself, if only to decree its inaccessibility; on the other side, although reality is purely *in itself*, it is bound to re-emerge as the negative of consciousness, if only to remain out of reach. For Hegel, this impasse is the final result not only of Platonism, Pyrrhonism and Kantianism, but of every critical project predicated on a rigid conception of negativity. This kind of critique cannot go any further because it is forced to return, again and again, to the dualism it seeks to overcome. Reduced to this pure negativity, consciousness 'must wait to see whether something new comes along and what it is, in order to throw it too into the same empty abyss.'²²

This impasse is also discussed near the end of the preface, where Hegel reiterates the opposition between a purely negative kind of scepticism, whose result is pure nothingness, and a positive or dialectical kind of scepticism, whose result is a determinate nothingness.²³ The first of these attitudes is linked to the procedure of 'argumentation' (*das rasonierende Verhalten*), and to what Hegel calls 'representational thinking' (*vorstellendes Denken*). The second one is described as a form of 'notional thinking' (*begreifendes Denken*).

The first attitude consists in a radical form of criticism, intent on refuting every idea or content it comes across. Whenever something is deemed real or true, the critic is quick to oppose it and refer back to the self: what is thought to be real is only so *for consciousness*, and hence unreal and untrue. But this attitude is once more dismissed as sterile and one-sided. This kind of reasoning, Hegel argues, 'imagines that by establishing the void it is always ahead of any insight rich in content',²⁴ but its argumentation is only 'a reflection into the empty "I", the vanity of its knowing.'²⁵ Its relationship with the content is still purely negative: since each new object is reduced to the same 'void', consciousness comes away empty-handed, and must look for a new object to negate.

²¹ PS, 50 / HW 3, 73

²² PS, 51 / HW 3, 74

²³ See PS, 36ff. / HW 3, 56ff.

²⁴ PS, 36 / HW 3, 57

²⁵ PS, 36 / HW 3, 56

To the ‘notional’ thinker, on the contrary, negativity is not abstract, but determinate. Every refutation is simultaneously a *positive* refutation because it *posits* something other than a complete nothingness. The negated content is not merely nullified, but also revealed as a new object, and as the origin of a new cognitive relationship. This kind of scepticism is the one employed throughout the *Phenomenology*. In order to emphasize the novelty of his method, Hegel goes on to spell out its actual stages, in a description that is evocative of the one offered in the introduction, but which is worth considering in detail.

The previous contrast between notion and reality, or truth and appearance, is now reframed by the grammatical contrast between subject and predicate. The difference between ‘representational’ scepticism and ‘notional’ scepticism hinges on the opposition between two different ways of understanding the subject-object relation, and hence on the logical structure of the predication process. In the first instance, consciousness sees itself as an inert subject, to which objects are appended in a contingent manner, as mere predicates. This mode of thinking can be illustrated by the following scheme:

Subject 1 → Predicate A

Subject 1 → Predicate B

Subject 1 → Predicate C

...

When something is predicated of a given subject without it being significantly altered, that can only mean one of two things: either the predicate in question is indeed purely contingent and does not change the subject it comes into contact with, or the predicate is already contained in the subject and amounts to the mere explicitation of an otherwise silent element, already part of the subject’s definition. The proposition *A is B*, for instance, does not reveal anything new about A insofar as a) A is just as well C, D or E, in which case its being B is not one of its defining features; or b) B is already part of A’s definition, in which case the proposition is but a development of the identity principle, $A = A$. In both cases, the predication does nothing more than confirm what was already known about the subject, without questioning its established identity. And that is why the procedure of argumentation appears so free and versatile. Its virtuosity seems boundless because its starting point and its finishing point are the same. There is no real predication, in the Hegelian sense, because the argumentative subject is really only affirming itself.

Notional thinking, on the contrary, triggers a dialectical inversion of the grammatical elements in play. The example provided by Hegel—the statement ‘the actual is the universal’²⁶—is the conceptual result of the *Phenomenology*’s first major *Aufhebung*, viz. the transition from the stage of sense-certainty to the stage of perception. In this statement, the predicate ‘universal’ is not merely discarded or

²⁶PS, 39 / HW 3, 60

swallowed up by the subject. The link between actuality and universality is neither accidental nor tautological, as in the previous model: universality is not a mere predicate of actuality, but the *notion*—or, as Hegel also puts it, the *substance*—underlying the whole relationship. Indeed, although actuality starts out by being actual, it loses its identity during the predication process. What comes to light when actuality is equated with universality is not that the actual is *also* the universal, or that the actual *happens to be* universal, but that the actual is, *after all*, the universal. Universality is not merely 'represented' (*vorgestellt*) by the subject, but grasped (*begriffen*) as its actual truth.

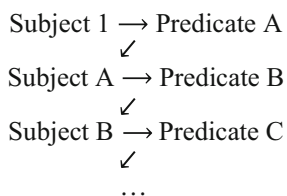
[Representational thinking] suffers, as we might put it, a counterthrust. Starting from the Subject as though this were a permanent ground, it finds that, since the Predicate is really the Substance, the Subject passed over into the Predicate, and, by this very fact, has been sublated; and, since in this way what seems to be the Predicate has become the whole and the independent mass, thinking cannot roam at will, but is impeded by this weight.²⁷

Unlike the procedure of argumentation, notional predication is not reversible. It destroys itself as it goes along.²⁸ Whereas in the previous model the subject extended peacefully to the predicate, recognizing itself therein, it is now absorbed by the predicate and forced to rethink its own identity. The subject no longer sees itself with its own eyes, but with the eyes of the predicate, and the latter becomes, therefore, the new subject:

Here thinking, instead of making progress in the transition from Subject to Predicate, in reality feels itself checked by the loss of the Subject. . . . And now, having returned into itself in the Predicate, instead of being in a position where it has freedom for argument, it is still absorbed in the content, or at least is faced with the demand that it should be.²⁹

The subject is no longer free because its notion is held hostage by the predicate. At the same time, however, the latter also loses its original freedom, i.e. the ability to be associated, generically, to other subjects. The predicate is now bound to the determinate negation of *this particular subject*: only as such is it in fact the new subject, as well as the source of a new proposition and a new dialectical inversion.

Once more, this movement can be illustrated by a simple scheme:



²⁷PS, 37 / HW 3, 58

²⁸Though it can be stated that the 'actual is the universal', it cannot be stated that 'the universal is a predicate of the actual'. The truth of the term 'actual' has migrated to the predicate, thereby transforming the entire relationship.

²⁹PS, 38 / HW 3, 59

The *Phenomenology* is modelled on this dynamic form of predication. It consists in a global system of subjects and predicates whose meaning is successively transferred to new predicates and new subjects. In other words, it consists in the repeated acknowledgement that in thinking or saying something ‘we meant something other than what we meant to mean; and this correction of our meaning compels our knowing to go back to the proposition, and understand it in some other way.’³⁰ This procedure has the following form: consciousness starts out by assuming that all knowledge is immediate knowledge. It begins by claiming that the ‘true is the actual’ (Subject 1 → Predicate A). However, by pursuing this claim and fleshing out its consequences, it realizes that it must turn back and correct its original proposition. The ‘actual’ is not what it seems, but its opposite, namely the ‘universal’ (Subject A → Predicate B). This new predicate becomes the new subject, but its meaning is once again called into question: the ‘universal’ is not what it seems, but the ‘infinite’ (Subject B → Predicate C); and the latter, in turn, will also prove inconsistent and in need of replacement. This cycle can only be broken once consciousness says what it actually means to say, and no longer needs to correct itself. This can only happen, according to Hegel, in the stage of science, or ‘absolute knowing’, where ‘*Subject* is just as much *Substance*’,³¹ and predication finally comes to an end.

In the preface, as in the introduction, Hegel insists that truth cannot be known immediately or directly, but only as a *result*. Immediate truths, because they are immediate, amount to bare assumptions, which can be challenged and replaced by other assumptions. Asserting a preposition by ‘adducing reasons for it, and in the same way refuting its opposite by reasons, is not the form in which truth can appear.’³² If the latter is ever to be reached, consciousness must rise above this arbitrariness, and it can only do so in an immanent way, through ‘the exposition of the untrue consciousness in its untruth.’³³ But although the acknowledgement that one’s assumptions are untrue is already a step further from believing them, it is still a long way from actual truth. What distinguishes Hegel’s critical approach from conventional scepticism is the idea that untruth is not only negative, but also

³⁰PS, 39 / HW 3, 60. This self-correcting procedure is one of the main causes of the frequent complaint regarding the unintelligibility of Hegel’s philosophy. In an essay on the clarity of Hegel’s writings, Adorno argued that this complaint is largely due to the un-Hegelian assumption that ‘[philosophical] terms must remain within the meaning once given them by definition’ (*Hegel*, 112). This assumption is un-Hegelian because it confines philosophical reason to the static poles of self-identity and difference, and regards intelligibility as a fixed measure of the identity of a given idea with itself. However, ‘once it is acknowledged that clarity and distinctness are not mere characteristics of what is given, and are not themselves given, one can no longer evaluate the worth of knowledge in terms of how clearly and unequivocally individual items of knowledge present themselves’ (ibid., 100). Rather, ‘only when the life of the thing expressed by the concept [*Begriff*] is compared with the meaning specified and when the old meaning is thereby dishonoured as invalid, is the other meaning constituted.’ (113). Michael Inwood makes a similar point in *A Hegel Dictionary*, 14.

³¹PS, 489 / HW 3, 587

³²PS, 28 / HW 3, 47

³³PS, 50 / HW 3, 73

positive—that it can be set in motion, as it were, and lead the way to its own supersession.

Untruth is not the end, but the beginning. However, given the *Phenomenology's* peculiar methodology, the success of Hegel's undertaking is conditional on at least three major requirements. First, if truth is incompatible with the arbitrariness of human assumptions, it can only be secured when the latter is fully overcome. For this to happen, consciousness must expose the untruth not only of this or that assumption or set of assumptions, but of *every single one of them*. It must traverse the entire realm of appearance and make sure that no assumption is left uncriticized. Otherwise, when it finally finds truth, it will have no way of knowing whether it is indeed so or just a new assumption in disguise.

It might be said, returning to Hegel's grammatical model, that the *Phenomenology*, in order to be successful, must cover the entire range of propositions that can be endorsed by consciousness. Only when every subject and every object have been confronted and sublated, and there is nothing left to 'mean', is the realm of untruth overcome. The 'recognition of the failure of all criteria is the negative side of an argument which leaves Hegel's dialectical holism as the positive side, the only alternative. Emergent holism is not one truth among others, but the only truth there can be.'³⁴

Second, the success of Hegel's project is also conditional on the ability to identify the correct outcome of each proposition and the correct sequence of sublations leading to the standpoint of science. For the realm of appearance to be overcome, consciousness must follow a specific guiding thread. If the predicative chain is broken, or any of its links overlooked, the whole progression may be compromised. As Hegel puts it, the road to science *is itself already science*: its specific direction is not contingent or arbitrary, but determined by the notion of truth itself. Accordingly, the *Phenomenology* does not display only the *experience* of consciousness, but also the *science of the experience* of consciousness. 'The length of this path has to be endured, because . . . each moment is necessary' and 'has to be *lingered* over.'³⁵

Third, and no less importantly, the success of this undertaking is also dependent on there being in fact a finishing line. Even if Hegel is right about the self-evolving nature of human reason, and even if the progression's different stages are indeed necessarily linked to each other, the original disagreement between subject and predicate might never be truly eliminated and the progression might go on

³⁴Williams, *Recognition*, 112. This issue leads back to the classical notion of equipollence. Following the sceptical tradition, Hegel argues that the truth of a given claim can only be established once all claims to the contrary have been refuted. Until then, truth and untruth are merely dogmatic labels, arbitrarily pinned on unverified beliefs. Unlike scepticism's equipollence, however, Hegel's equipollence is not a cul-de-sac, but a principle of motion. He aims to overcome the sceptical impasse by refuting *all* of consciousness' positive claims about reality, thereby reaching the *negative of untruth*, i.e. truth or science. See Forster, *Hegel's Idea of a Phenomenology of Spirit*, 164–174; Röttges, *Der Begriff der Methode in der Philosophie Hegels*, 54–62; Trisokkas, *Pyrrhonian Scepticism and Hegel's Theory of Judgement*, 71–92.

³⁵PS, 17 / HW 3, 33

indefinitely. This would mean, once again, the failure of Hegel's project: if the progression had no ending, all of its stages, however sophisticated, might be refuted, and science would remain out of reach. Aware of this possibility, Hegel argues that the 'the *goal* is as necessarily fixed for consciousness as the serial progression',³⁶ and that 'the interconnection of the forms of the unreal consciousness will by itself bring to pass the *completion* of the series.'³⁷ Furthermore, he claims that the progress towards this final conclusion is unhalting 'and short of it no satisfaction is to be found at any of the stations on the way.'³⁸

These and similar claims are grounded, once more, in the idea of totality. The *Phenomenology's* stages are not simply untrue standpoints wrongly assumed to be true, but also moments of a wider, but invisible or inaccessible whole. The progression is not merely about truth and error, but also about a gradual movement of expansion, whereby consciousness is led not only to rephrase what it originally 'meant', but also to redimension the scope of its sentences. Again, however, this movement must not be construed in a quantitative way, as the gradual enlargement of the same cognitive standpoint, or the simple accumulation of different cognitive fragments. Each fragment is perceived, at first, as the whole world, and its refutation amounts therefore to a global refutation, leading to the embracement of a new cognitive totality. But although this new standpoint is just as global as the first one, and just as untrue, it offers nonetheless a wider (or less limited) glimpse into the progression's final standpoint.

This idea can be clarified with the aid of Leibniz and his own characterization of human cognition. In Leibniz's epistemology, cognition is also grounded in a global account of reality. However, each individual consciousness is confined to an individual cognitive perspective: each monad is a single *and yet* a global standpoint, a fragment *and yet* the whole universe. This paradox is illustrated by the image of an observer looking at the same city from different angles: depending on his or her position, the observer sees the same elements combined in different ways.³⁹ This idea is reiterated in the *Phenomenology*, where each new standpoint is also a situated outlook on the whole of reality. For Hegel, however, this situation is not definitive. Whereas Leibniz's description is focused on the simultaneous diversity of different individual perspectives, Hegel's is focused on the self-evolving diversity of the same cognitive perspective. Instead of different observers placed in different positions, each confined to a single version of reality, the *Phenomenology* documents the movement of the same observer as he or she adopts different cognitive angles.

In Hegel's account, these angles are not equivalent. Although they are all equally one-sided, each builds on the information provided by the previous ones, like a cubist composition with increasingly faceted shapes. As each new angle is criticized and sublated, it 'ranges itself as a moment having its own place in the whole'⁴⁰ and

³⁶PS, 51 / HW 3, 74

³⁷PS, 50 / HW 3, 73

³⁸PS, 51 / HW 3, 74

³⁹See Leibniz, *Die philosophischen Schriften*, vol. 1, 10, 69; vol. 4, 434; vol. 6, 616.

⁴⁰PS, 32 / HW 3, 52

becomes part of the cognitive ground on which a new standpoint is made to rest. Due to the nature of this method, truth can only be attained once all different angles have been viewed and criticized, which brings us back to the third requirement listed above, and to another major difference between the Hegelian and the Leibnizian models. Unlike Leibniz's spatial metaphor, where the same city is viewed from an infinity of different perspectives, Hegel draws up a finite progression, comprising a limited number of stages. He argues that the whole underlying every cognitive perspective can be divided into a specific number of parts, or sets of parts, which sums up all of reason's cognitive attitudes. This does not mean, of course, that consciousness is reduced to a finite number of thoughts or judgements: to be sure, each phenomenological stage allows for an infinite number of instantiations, prompted by the infinite variety of human experience.⁴¹ But all of one's thoughts fall under a finite number of 'shapes of consciousness', whose sublation implies the automatic sublation of all of their subproducts and ramifications.

6.3 Further Hermeneutical Issues

The *Phenomenology of Spirit* documents the long transition from a natural or uneducated mode of cognition to a scientific or educated one. Having outlined, in the previous sections, Hegel's methodological strategy and the main requisites for the successful completion of his project, I will now return to the beginning and focus on the ignorance he places at the start of the progression.

If natural or phenomenal consciousness (*das erscheinende Bewußtsein*) corresponds to our usual cognitive perspective, it would seem natural to interpret Hegel's progression as an actual history of human reason, where the different phenomenological stages would amount to actual cognitive strata, leading from the most primitive to the most sophisticated mode of cognition. In the stage of sense-certainty, knowledge would be a mere apprehending, focused on the sheer being of things.⁴² In the stage of perception, it would learn to distinguish identity from difference, truth from falsehood. With perception, it would formulate laws, establish causal relations, make logical inferences, etc. Its knowledge would be gradually perfected until all its initial limitations were overcome.

The problem with this kind of reading is all too apparent. Although our usual cognitive perspective is far from perfect, it is certainly not as limited as the stage of sense-certainty. In its natural or uneducated form, it is already capable of much more than simply apprehending reality: it is able to recollect past events, to anticipate future ones, to confront different ideas and to form highly complex judgements. As we have seen in connection with the *Philebus* and Plato's cognitive scale, human

⁴¹ See PR, 205 / HW 7, 370: 'Completeness means the exhaustive collection of every single thing pertaining to a given field, and no science or branch of knowledge can be complete in this sense.'

⁴² PS, 58

cognition is intrinsically irreducible to the stage of sense perception. The human mind is incapable of envisaging a purely sensuous account of reality and can only consider such a possibility in a negative way, as the inner contraction of a more complex cognitive regime. And the same holds for the stage of recollection, viz. for a purely mnemonic account of reality. The native ground of human cognition is the stage of *δοξάζειν*, already bound by normativity and predication.⁴³

In the beginning of the *Phenomenology*'s first chapter, Hegel claims that the kind of knowledge 'which is *at the start* or is *immediately* our object' is sense-certainty.⁴⁴ But these adverbial expressions must not be interpreted in a literal manner, as direct qualifications of our actual cognitive standpoint. The progression does not start with the most immediate form of knowledge there is, but with *one's usual conception* of the most immediate form of knowledge there is. In other words, although natural consciousness is aware that its actual knowledge is much more 'advanced' than sense-certainty, it tends to recognize the latter as the foundation of human cognition, and its own knowledge as a direct development thereof. What is at stake in the beginning is thus a *retrospective* diagnosis: starting from its actual standpoint, consciousness retraces its own cognitive history all the way to the very beginning. And the very beginning is sense-certainty.

Hegel's progression does not consist in a series of actual cognitive stages, but in a series of cognitive models, or 'shapes of consciousness'. It sets out to offer a complete anatomical and morphological description of human cognition, starting with its most basic elements. Unlike other similar descriptions, however, this one is not univocal, nor offered by a neutral or disinterested observer: the different models taken up and criticized throughout the *Phenomenology* are forged by consciousness itself to account for its own constitution.

The acknowledgement of this aspect is paramount, in my view, to a proper understanding of Hegel's project. It can help explain not only the disparity between the progression's initial standpoint and our usual standpoint, but also the apparent incongruity of some of Hegel's thematic and historical transitions. Since the logic binding the different stages together is an immanent logic, determined by consciousness' self-understanding, the progression's guiding thread is not factual or historical: it does not reflect the sequence of cognitive attitudes adopted throughout human life, nor the sequence of philosophical doctrines endorsed throughout human history, but a long chain of cognitive self-portraits, as it were. Throughout the *Phenomenology*, consciousness is led to examine different misconceptions about its own cognitive nature, in search of 'the right sort of self-understanding'.⁴⁵

If the phenomenological progression is a retrospective progression in the aforementioned sense, its starting point is not an actual cognitive stage, but a specific cognitive regime. By placing sense-certainty at the basis of human cognition, consciousness declares its own knowledge to be a development of sense perception,

⁴³ See Sect. 2.1.

⁴⁴ PS, 58 / HW 3, 82. Emphases added.

⁴⁵ Patten, *Hegel's Idea of Freedom*, 124

that is, a highly complex web of simple sensory units, brought together to form different empirical constructs. This initial stage corresponds, therefore, to a simple form of empiricism, which proposes to account for highly developed forms of meaning while remaining faithful to the simplest of cognitive formulas. Following this logic, one can even envisage different versions of the progression's subsequent stages based on sense-certainty: an empirically-based self-consciousness, an empirically-based reason, an empirically-based morality or an empirically-based religion. Unlike the cognitive models introduced in the stages of self-consciousness, reason, morality or religion, all of these versions would presuppose the atomistic understanding of reality characteristic of sense-certainty.⁴⁶

When the stage of sense-certainty is sublated, consciousness realizes that this simple form of empiricism is not the real source of human cognition. More precisely, it realizes that the complexity of its actual mode of cognition cannot be reduced to a mosaic of sensuous units, and hence that the progression's real starting point must lie elsewhere. As sense-certainty gives way to perception, consciousness puts forward a new idea of what the simplest form of cognition might be. Instead of a simple empirical atomism, it embraces a more complex form of atomism, made up of 'things' and their properties. Once again, however, it realizes that its actual knowledge is irreducible to this new cognitive framework: consciousness is already a mode of 'understanding'—and then a variation of self-consciousness, and afterwards a modality of reason, etc. With each new transition, the progression's real starting point is pushed forward, until the very last stage. Since the entire *Phenomenology* is about finding a true cognitive foundation, unsusceptible of being denounced as yet another false start, and since this discovery does not occur in any of the progression's intermediate stages, Hegel's point seems to be that the true foundation of human cognition is nothing less than absolute knowing, viz. the dialectical synthesis of notion and object, essence and appearance, for-itself and in-itself. In other words, his point seems to be that our usual perspective is *already* scientific—a scientific perspective *in germine*, unaware of its true nature and forced to endure a long journey of self-discovery.

But none of this is known to the consciousness that embarks, for the first time, in the *Phenomenology*. The realization that the first cognitive model is untenable leads the phenomenological traveller to look for a new one; and when the latter is refuted, a third model is brought in to replace it. Each new solution is construed as a new alternative, disconnected from the previous ones. The immanent necessity binding the different stages is not yet apparent and the progression is not yet lived as a

⁴⁶In the *Encyclopaedia*, when discussing the immanent nature of the progression leading from 'subjective spirit' to 'absolute spirit', Hegel observes that 'the determinations and grades of Spirit . . . are essentially only moments, states, determinations in the higher stages of development. Consequently, a lower and more abstract determination of the mind [*Geist*] reveals the presence in it, even empirically, of a higher phase. In sensation, for example, we can find all the higher phases of the mind as its content or determinacy. And so sensation, which is just an abstract form, may to the superficial glance seem to be the essential seat and even the root of that higher content, the religious, the ethical, and so on; and it may seem necessary to consider the determinations of this content as particular species of sensation.' (PM, 8f. / HW 10, 16f.)

progression. Consciousness only sees the negative side of experience and not yet its positive value.

But although this uneducated standpoint is the *Phenomenology*'s main protagonist, it is not the progression's only standpoint. In the introduction, when discussing the work's methodology, Hegel distinguishes two different angles from which to look at the progression. The first reveals a pathway *for consciousness*, corresponding to the negative attitude just mentioned. The second reveals a pathway *for us*, philosophers or 'phenomenologists'. Natural consciousness is unaware of the progression's immanent logic, but *we*, philosophical observers, know each stage to be the necessary result of the previous one, and a piece within a wider dialectical puzzle. For *us*, each *Aufhebung* is both a suppression, a preservation, and a transition to a higher and more comprehensive standpoint.

This hermeneutical contrast is summed up as follows:

[With each transition,] a new pattern of consciousness comes on the scene . . . , for which the essence is something different from what it was at the preceding stage. It is this fact that guides the entire series of the patterns of consciousness in their necessary sequence. But it is just this necessity itself, or the *origination* of the new object, that presents itself to consciousness without its understanding how this happens, which proceeds for us, as it were, behind the back of consciousness. . . . *For it*, what has thus arisen exists only as an object; *for us*, it appears at the same time as movement and a process of becoming.⁴⁷

Caught up in experience, consciousness is unaware of its underlying logic. All it sees is a succession of preformed objects, whose origination can only be viewed from a wider hermeneutical angle, for which the movement itself takes precedence over the different stopping points. Such an angle—*our* angle—is no longer dominated by a simple or tautological conception of identity. It knows each phenomenological transition to be both a negation and a position, the collapse of a cognitive world and the emergence of a new one. Therefore, it regards each new object not merely as something that *happened to emerge*, but as something that was *bound to emerge* and is *bound to be overcome* in a determinate way. This is why Hegel claims, in the introduction, that only *for us* is 'the way to science' 'already 'science'—or that only from *our* point of view is the 'experience of consciousness' simultaneously the 'science of the experience of consciousness'.⁴⁸

Given this contrast, the *Phenomenology* is often divided into two textual layers. The first, *empirical* one, is the layer of action, so to speak, and usually occupies the central section of each chapter. It comprises 'the passages where consciousness is urged, as it were, to put itself in motion, . . . in order to adapt its vision of things to the things that are actually brought into view'.⁴⁹ The second, *scientific* layer, usually limited to the moments of transition, provides an interpretation of what goes on in the first layer. It comprises 'the developments during which consciousness takes a break,

⁴⁷PS, 56 / HW 3, 76

⁴⁸Ibid.

⁴⁹Labarrière, *La Phénoménologie de l'Esprit de Hegel*, 37

so to speak, while the philosopher steps to the foreground, in order to account for the conceptual continuity that makes each experience the *necessary* product of the preceding one.⁵⁰ This second layer is strictly exegetical: it does not interfere with the progression's actual unfolding. Since the critique pursued throughout the *Phenomenology* is immanent and self-moving, all that is left to the phenomenological observer is 'simply to look on'.⁵¹ But this passive role is just as essential as the active role played by consciousness: only through this added interpretation, 'contributed by us', is 'the succession of experiences through which consciousness passes [raised] into a scientific progression'.⁵²

The contrast between these two layers is indicative of a wider structural distinction, which poses an additional hermeneutical challenge. If the only 'way to Science' is a phenomenology of spirit, the 'scientific' sections of the text cannot be fully understood by ordinary consciousness. Only a standpoint that has already gone through the progression and attained its final stage can grasp the true meaning of these passages. This conclusion suggests that Hegel's progression is not a simple, but a double progression, which requires a double interpretation: after the elevation of natural consciousness to the stage of science, the resulting standpoint must go back to the beginning and acknowledge the immanent necessity underlying its transformation. In this second reading, 'Science sets forth this formative process in all its detail and necessity, exposing the mature configuration of everything which has already been reduced to a moment and property of Spirit.'⁵³

Yet this kind of approach is contradicted by at least three main aspects. First, although the progression's initial standpoint corresponds to consciousness' usual standpoint—or rather, to its usual conception of the most simple form of knowledge—the phenomenologist's remarks are not entirely unintelligible to the reader that first takes up the *Phenomenology*. To be sure, the logical arguments and connections 'contributed by us' are not immediately clear. They fall outside ordinary consciousness' usual cognitive horizon and require the adaptation to a new language and a new set of concerns. Nevertheless, their meaning is taken into account from the very beginning and becomes increasingly familiar. Second, the alternation between the progression's empirical and scientific layers does not follow a fixed pattern. Whereas the first chapters are structured in a relatively similar way, in line with the introduction's methodological description, the more advanced stages are not as

⁵⁰ Ibid.

⁵¹ PS, 54 / HW 3, 77

⁵² PS, 55 / HW 3, 79

⁵³ PS, 17 / HW 3, 33. Returning to the parallel between the *Phenomenology* and the allegory of the cave, this retrospective standpoint can be compared to the liberated prisoner's standpoint as he returns to the cave and attempts to release the other prisoners. Having regained his initial position, his second ascent is very different from the first: he already knows the shadows to be mere shadows, the cave to be a mere cave and the various forms of illusion found inside and outside the cave to be the intermediate stages of a long cognitive journey.

uniform. The balance between the two layers is at times highly uneven, and their length and importance very unequal. Third, the ‘double reading theory’ presupposes the (relative) independence of the empirical and the scientific layers. It presupposes the possibility of a first reading exclusively focused on the first layer, free from the retrospective element introduced by the second one. But in most cases this separation is very difficult, if not impossible. The phenomenologist’s considerations are frequently woven into the main text and cannot be eliminated without sacrificing its intelligibility. Moreover, the very frontier separating the two registers is often hard to determine and grows increasingly blurred as the work progresses.⁵⁴

For all of these reasons, the *Phenomenology* is best construed as a single, unified movement, made up of different but simultaneous hermeneutical layers. The central distinction between phenomenal and scientific knowledge, reiterated and transformed at each new stage, is complemented by the secondary, but no less important distinction between consciousness’ and the philosophical narrator’s understandings of the progression’s immanent logic. Throughout the text, the distance separating these two standpoints is increasingly reduced, and finally eliminated. The standpoint of ‘absolute knowing’ signals therefore not only the coincidence between reality *in itself* and reality *for consciousness*, but also the fact that this coincidence is revealed both *for us* and *for consciousness*. In other words, the elevation to the standpoint of science entails the simultaneous acquisition, on the part of consciousness, of a clear understanding of the entire progression.

⁵⁴ See Labarrière, *La Phénoménologie de l’Esprit de Hegel*, 35–38.

Chapter 7

Knowledge, Power and Freedom



Abstract Looking at the *Phenomenology of Spirit's* Consciousness and Self-consciousness sections, I discuss the three main antitheses that fuel Hegel's work: the oppositions between subjectivity and objectivity, between subjectivity and intersubjectivity, and between a secular and a transcendent mode of cognition. While Sect. 7.1 examines the progression's initial stages, from the naïve empiricism of 'sense-certainty' to the more sophisticated, but equally abstract standpoints of perception and the understanding, Sect. 7.2 analyses the emergence of self-consciousness and the dialectics of desire, life and recognition. I argue that Hegel's conception of recognition amounts to an important redefinition of the relationship between freedom and education, and the first clear break with the critical framework inherited from Plato. Afterwards, Sect. 7.3 revisits Hegel's master-slave dialectic and the transition to stoicism and scepticism. Here, again, the subjection endured by the slave is revealed as a cognitive imprisonment, to be overturned by a philosophical reform of human cognition. Finally, Sect. 7.1 looks at the last stage of the Self-consciousness section. Whereas previously the progression was confined to a horizontal plane, inhabited by human subjects and natural objects, the 'unhappy consciousness transfers' the debate to a vertical plane, introducing the contrast between a human and a divine standpoint.

Keywords Hegel · *Phenomenology of Spirit* · Consciousness · Self-consciousness · Dialectic · Freedom · Master · Slave

The *Phenomenology's* first great section, dedicated to 'consciousness' in general, offers a series of variations on one central theme: the automatic and widespread belief in the intrinsic independence of the objective world. According to Hegel, ordinary reason is defined by the assumption that reality exists in its own right, regardless of one's knowledge of it. Things are what they are and can either be known or ignored, perceived in a clearer or less clear way. The world is a show in which the subject is a spectator, confronted with ready-made objects and events. This does not mean, of course, that reality cannot be transformed. The subject is a

spectator, but also an actor, whose actions have real effects and consequences. Yet all of those actions presuppose the acknowledgement of a fixed objective setting, in which subjectivity is primarily grounded.

In the beginning, this assumption is taken up in its most primitive form. The object of consciousness is one that simply *is*, ‘regardless of whether it is known or not’,¹ and remains so ‘even if it is not known’.² But although consciousness is ready to buy into this version of things, a closer look at objectivity will bring to light its various contradictions. By attempting to capture the object in its independence and immediacy, consciousness will realize that the object is neither independent nor immediate: instead of simply being there, objective reality is already implicated in a dialectical relationship with the knowing subject. Its identity is not simple, but mediated, and its objectivity is stained with subjectivity.

This discovery is the motor of the first three chapters. The realization that objectivity is not objective prompts the abandonment of the progression’s first stage and the adoption of a new one. But as consciousness moves from one standpoint to another, the subjective contamination of objectivity increases. Finally, at the end of the first section, subjectivity takes over completely. The world is revealed as the very opposite of a fixed objective setting and consciousness ends up becoming its own object. This transformation inaugurates the *Phenomenology*’s second great section, dedicated to ‘self-consciousness’. The genesis and development of this new phenomenological shape will be briefly examined in the following pages.

7.1 Selfhood and Otherness

Initially, cognition is nothing but ‘sense-certainty’ and its object sensuous reality as a *factum brutum*, prior to any intellectual or categorial distinction. On the subjective side of the relationship, consciousness amounts to a transparent filter, which reveals the world as it really is. On the objective side, reality is made up of simple, concrete sensuous units, which reveal themselves immediately to consciousness. On closer inspection, however, the first contradictions start to emerge. Reality may well be simple and concrete, but this simplicity and this concreteness are not within the reach of sense-certainty. Consciousness cannot conceive of such a simple world because the very concepts it uses to refer to it—‘this’, ‘here’ and ‘now’—are already universal and abstract.³ Their meaning is not immediate or self-contained, but

¹PS, 59 / HW 3, 84

²Ibid.

³To *abstract* (*abs-trahere*, ἀ-φαίρειν) means to ‘draw away’ or to ‘remove’ something from something else. For Hegel, something is *abstract* when its meaning is unduly isolated and posited as self-sufficient. Something is *concrete*, on the other hand, when it is grasped as part of a wider dialectical whole. See LHP, vol. 1, 23–27 / HW 18, 42–46.

inherently transitive: it points to something *other* than themselves, without which it would not be intelligible. Although ‘this’, ‘here’ and ‘now’ are used to refer to simple things, they are in fact undifferentiated conceptual operators, which can be applied to an infinity of objective contents.

This contradiction prompts the endorsement of a new cognitive thesis: simplicity and concreteness do not reside in reality, but in consciousness: it is *for consciousness* that a tree is ‘this tree’ and not another, and that the present moment is ‘this now’ and not another. But like the ‘here’ and ‘now’ of the previous stage, the ‘I’ of whom these claims are predicated—the ‘I’ which guarantees the singularity of each object—is itself revealed as an undifferentiated operator. Since the ‘I’ is always an ‘I that x’, or an ‘I that y’, it is never a simple or self-sufficient determination, but a universal and abstract one, which can be applied to an infinity of different contents.

The truth of the singularity envisaged by consciousness cannot be found either in the subject or in the object. In a last attempt to preserve its initial standpoint, consciousness turns to the relationship itself: if both the subject and the object, when considered separately, turn out to be universal, the loss of singularity can only arise from the very act of isolating them. To correct this, consciousness must focus on the two elements simultaneously: viz. an individual ‘I’—‘this I’—facing a ‘here’ and a ‘now’—‘this here’ and ‘this now’—and preventing any form of otherness from undoing the immediacy of this knowledge. However, in the very act of singling out a specific temporal moment, consciousness loses what it means to grasp. Every ‘now’, in being pointed out, ceases to be what it was: it is immediately replaced by a new ‘now’, doomed in turn to disappear. Likewise, every ‘here’, in being indicated, is revealed as part of a complex system of spatial relations: it is simultaneously ‘above’ something and ‘below’ something else, a ‘right side’ and a ‘left side’, an ‘inside’ and an ‘outside’. Single spatial moments, like single temporal moments, are pure abstractions, artificially removed from a universal continuum where every element is intrinsically dependent on all others.

Just like the subject and the object of sense-certainty, the whole relationship rests on universal, abstract operators, and this new acknowledgement leads consciousness to ‘go back to the [initial] proposition, and understand it in some other way.’⁴ The simplest form of cognition is thus transferred from the stage of sense-certainty to the stage of ‘perception’ (*Wahrnehmung*). Whilst previously reality was a universal aggregate of sensuous units, it is now conceived as a system of ‘things’. And the experience of this new phenomenological stage also follows a triadic scheme. First, unlike a single sensuous unit, each thing is both simple—it is *this* thing as opposed to others—and universal—it partakes of a series of universal properties, shared by many other things. The first of these aspects is the thing’s *exclusive* unity. It is that in virtue of which a given thing is perceived as unique and self-identical: as an object whose colour, for instance, is *its particular colour* and not a generic colour, and whose shape is *its particular shape* and not a generic shape. The second aspect is the thing’s *inclusive* unity, whereby it is perceived as an assemblage of universal

⁴PS, 39 / HW 3, 59

properties: as an object whose colour is an instantiation of a generic colour, whose shape is an instantiation of a generic shape, etc. As consciousness explores these two aspects, it comes to recognize their incompatibility. If something is unique, the universality of its properties cannot be maintained. The thing's predicates must be subsumed under a truly exclusive and self-identical unity. Conversely, if something is varied, its uniqueness cannot be maintained. It ends up being lost, reduced to a more or less complex aggregate of universal properties. The 'thing' is thus an inherently contradictory object caught between absolute selfsameness (nominalism) and absolute otherness (indetermination).⁵

This first contradiction is followed by a subjective moment, where consciousness takes responsibility for the thing's inconsistencies and attempts to save the perceptual model. However, in contrast to what happened in the first chapter, consciousness does not intervene as the essential moment of the relationship, but as the source of its untruth. Instead of perceiving the thing as it actually is, perceptual cognition adds a reflective element to the relationship, which distorts the thing's true image. The contradiction between universality and singularity is thus not inherent to the thing itself, but the result of an observational defect, which must now be corrected.

This correction is itself divided into two symmetrical moments, which correspond to the second and third parts of Hegel's argument. To begin with, consciousness accepts the thing's exclusive unity and regards its different properties as subjective additions. This means, for instance, that a grain of salt is not white in itself, but only in the eyes of the subject; that it is not cold in itself, but only because it is perceived as such; that it is not tart in itself, but only when the subject tastes it, etc. Things in themselves are only varied because the subject's corporeal senses break them down into different aspects and perceive them in different ways. But this attempt to preserve the thing's simplicity is unsuccessful, as it fails to acknowledge that simplicity itself is always a *specific* or *determinate* simplicity. There is no identity, however simple, without determination: for something to be what it is, it must be different from everything else, and this difference can only be found in its properties.

Consciousness then reverses its initial position, accepts the thing's intrinsic multiplicity and recognizes itself as the source of the thing's unity. The grain of salt is indeed white *and* cold *and* tart—but *insofar* as it is white it is not cold, and *insofar* as it is cold it is not tart. It is the perceiving subject who, in order to escape the contradiction, isolates each of the properties and sets them up as independent categories. But this hypothesis is also problematic. If the link binding each thing to its properties is merely imaginary, the thing's identity is itself imaginary. Things are reduced to mere 'enclosing surfaces', which can hold an endless variety of contents.

⁵Hegel's argument leads back to Plato's μέθεξις. Though the definition of a thing is necessarily supported by a series of εἶδη, or γένη (redness, roundness, bitterness, etc.), each εἶδος or γένος can only be defined by providing examples of its individual occurrences (a red apple, a round apple, a bitter apple, etc.). See Sect. 4.1.

By focusing alternately on the objective and the subjective poles of the relationship, consciousness attempted to contain the contradiction between identity and difference. But experience has shown that this contradiction is embedded in the very heart of cognition—that it ‘circulate[s] everywhere in every thought that is uttered’⁶ and cannot be cast aside. While the perceptual model is still guided by a tautological mode of cognition, in which self-identity is reduced to selfsameness, the *Phenomenology*’s third chapter will introduce a different cognitive framework, where selfsameness and otherness are irrevocably intertwined.

Although the singularity pursued in the stage of sense-certainty passed over into universality, the latter is still, as Hegel puts it, a *conditioned* universality. In the stage of perception, the main contradiction of sense-certainty was not removed, but merely transferred to the more complex domain of thinghood. The thing is universal and *at the same time* determinate; it is one and *at the same time* varied. This paradox can only be overcome once the two competing terms cease to be posited alongside each other. The truth of perception is not disjunctive, but dialectical: it is not one of these two alternatives, nor the two of them together, but the very movement that separates and unites them. It is the dialectical transit whereby identity turns into otherness and vice versa.

With this discovery, the composite universality of perception is replaced by a genuine or *unconditioned* universality.⁷ While the two previous models attempted to maintain the subject’s and/or the object’s independence, the ‘understanding’ (*Verstand*) favours the relationship itself over its terms. This conceptual reversal introduces a new kind of object, ‘force’ (*Kraft*), which can no longer be thought of in atomistic or dichotomic terms.

The dialectical impasse which led to the implosion of thinghood is now recast as the internal duplicity of the same object: on the one hand, the movement whereby force loses its unity and is dissolved into the multiplicity of its properties is force’s ‘expression’; on the other hand, the counter-movement whereby this plurality is reunified is ‘force proper’, or force ‘*driven back* into itself’.⁸ Unlike the previous model, however, the interaction between these two poles is not added to the poles themselves. It is the primary element of the relationship.

To understand Hegel’s argument, it is useful to bear in mind the historical references scattered throughout the text. The transition from the dialectic of thinghood to the dialectic of force entails a departure from the world of common sense and the embracement of a strictly metaphysical account of reality. Forces are the operating units of a new approach to the natural world, whose historical emergence can be traced back to the seventeenth century, and particularly to the Leibnizian critique of Descartes’ mechanical physics. While the Cartesian model reduces objectivity to extension, and defines every object in terms of shape, size and

⁶*Phi* 15d4–5

⁷*Unbedingte Allgemeinheit*. Hegel plays with the radical ‘-ding’ and its negation. Unconditioned universality is the negation of *Dingheit*, i.e. ‘thinghood’.

⁸PS, 81 / HW 3, 110

motion, Leibniz argues that the world cannot be grasped in this way. Shape and size are intrinsically relative qualities, and hence an insufficient basis in which to ground identity and difference. Accordingly, reality cannot be reduced to a multiplicity of physical or geometric atoms, brought together as a collection. It consists, rather, in a series of self-determining units—‘formal atoms’, or monads—joined together by a complex system of forces.⁹

Like Leibniz, Hegel replaces a world made up of things with a thoroughly dynamic world, where all events and transformations are the expression of different forces. Moreover, he also insists on the *systematic* nature of this model. Whereas ordinary consciousness tends to construe forces as regional or isolated events, added to an originally static reality, the understanding takes over the entire realm of objectivity and envisages a world in which all traces of atomism have been eliminated.

The transition from the consideration of a single force to the description of a global system of forces is not accidental, but follows from the very notion of force. The starting point of Hegel’s argument is the movement whereby a given force expresses itself and retreats back into itself. These two moments are part of the same force, but the passage from one to the other seems to imply the intervention of an external element. The unexpressed force expresses itself because it is *solicited* to do so by something other than itself; likewise, the expressed force retreats back into itself due to a similar solicitation. But when this external motor is closely examined, its structure is found to be identical with that of force. Like the latter, it also alternates between an unexpressed form (as it solicits the force to express itself) and an expressed form (as it solicits the force to retreat back into itself). The acknowledgment of this symmetry leads the first force to break up into two forces, which both express (and solicit) and are expressed by (and solicited by) the other.

But these two forces do not exist in isolation. They are part of a wider ‘play of forces’, whose scope extends, in the course of experience, to the whole world. The identity of each force is expressed by another force, and the latter’s identity is expressed in turn by another. What emerges with this process is a truly systematic mode of identity, where everything is defined by its degree of interference with everything else.¹⁰ There is no longer any fixed point or pole, no independent or self-sufficient source of meaning. Identity is henceforth ‘a pure being-positing by another’.¹¹

As forces merge into each other, however, their reality is lost. Their essence now consists ‘simply and solely in this, that each *is* solely through the other, and what each thus is it immediately no longer is, since it *is* the other’.¹² The experience of force leads to its disappearance (*Verschwinden*): force evaporates into otherness, and

⁹See *Die philosophischen Schriften*, vol. 4, 444.

¹⁰In Leibniz’s vocabulary, this systematic identity is determined by specific forms of *entr’expression* and *entr’empêchement*. See *Die philosophischen Schriften*, vol. 4, 440.

¹¹PS, 85 / HW 3, 114. Translation modified.

¹²PS, 86 / HW 3, 115

what remains is merely the ‘notion’ or the ‘thought’ of it. This new acknowledgement is very important and lays the ground for all the chapter’s subsequent developments. The truth of the understanding is no longer the unconditioned universal envisaged as an object (i.e. force), but the very *idea* or *notion* of universality.

As in the previous stages, the objective moment of experience gives way to a subjective moment, where the object is no longer meaningful in itself, but only for consciousness. With the vanishing of force, the understanding loses its object, but this void cannot be filled in a simple or direct manner, by bringing in a new object or a new version of reality. This new transition no longer allows for any of that: having posited and tested different kinds of objects, consciousness was led from contradiction to contradiction, until the elimination of objectivity itself. What emerges with the disappearance of force is thus the acknowledgement that objectivity—all objectivity—is unreal or illusory. The reality cognized by consciousness is the mere thought of reality, and not reality itself. It is the *appearance* of reality, devoid of objective truth.

Since the understanding is now reduced to this simple acknowledgement, its new object can only be envisaged in a negative way, as that which lies *beyond* appearance. This negative region, placed outside the reach of the understanding, is ‘the true background of things.’¹³ Yet precisely because it is purely negative, it can only be conceived through what it is *not*, namely appearance. Therefore, the new object of the understanding is really a double object. It is ‘the syllogism which has for its extremes the inner being of Things and the Understanding, and for its middle term, appearance.’¹⁴

This syllogism is the one I have previously summed up with the equation $R + C = R'$, and its different variations. It is the basic formula of Kant’s critical philosophy, which Hegel now brings to the fore. As the understanding decrees the globally apparent nature of phenomenal reality and transfers truth to an ‘inner world’, it replicates the main dualism underlying Kant’s critique of the understanding. In both cases, consciousness is confined to a global cognitive perspective and bound to cognize things through the same transcendental lens. It is unable to step outside itself, as it were, and peak at the world’s true essence.

Unlike Kant, however, Hegel aims to show that this standstill is not definitive. ‘Certainly’, he claims, ‘we have no knowledge of this inner world as it is here in its immediacy; but not because Reason is too short-sighted or is limited, . . . but because of the simple nature of the matter in hand, that is to say, because in the *void* nothing is known, or . . . because the inner world is determined as the *beyond* of consciousness.’¹⁵ We find here the same objection raised in the introduction: critical philosophy is still grounded in an abstract conception of cognition; by separating thought from reality and building its whole argument upon this distinction it takes for granted the very obstacle it aims to criticize.

¹³PS, 86 / HW 3, 116

¹⁴PS, 88 / HW 3, 117

¹⁵PS, 88 / HW 3, 117f.

The remainder of the chapter is devoted precisely to overcoming this obstacle. Initially, since the 'inner world' is defined as the opposite of appearance, it is still an empty world. However, it will gradually 'fill itself out for the understanding'¹⁶ and reunite with consciousness. Hegel's description of this process is extremely nuanced, and can only be presented here in outline.

First, consciousness envisages the 'inner world' as the *law* that governs appearance. Although the phenomenal world is made of difference and change, the latter are not arbitrary. They are the unstable image of a stable realm of laws. But if the understanding manages hereby to join both terms of the syllogism, this realm of laws does not yet account for the entire realm of appearance. The phenomenal world retains a quantum of contingency that derives from the ambiguity of the very definition of law: though only a universal law can ensure the unity of the phenomenal world, its universality is too broad to determine anything specific; conversely, a law that seeks to translate the specific nature of phenomena splits up into a myriad of particular laws, contingently related to each other. At stake is the opposition between amplitude and determinacy: the broader a law, the less determinate its content, and vice-versa. The law by which a stone falls and the law by which heavenly bodies move can be joined under the same general law, namely universal attraction. But since the latter applies to everything, its content is hardly determinate. It does not express the two previous laws, but merely the *pure notion of law*, i.e. the generic declaration that 'all reality is conformable to law'. The discovery of this pure notion leads to a transformation of the dialectic's initial terms: instead of the contrast between a general law and a set of particular laws, consciousness focuses on the contrast between the different laws and their *notion*, or inner necessity. In the case of universal attraction, the law's notion is the simple factor 'gravity', but the law of gravity says something more specific, namely that two bodies are attracted to one another by means of a force that is directly proportional to their mass and inversely proportional to the distance between them. Likewise, 'electricity' corresponds to a simple notion, but the law of electricity postulates the attraction between a positive and a negative electric field. In both examples, however, the law does not follow necessarily from its notion. There is no real necessity because nothing in the notion anticipates the specific form assumed by the law: there is no logical reason why gravity should stipulate this specific relation and not another, or why electricity should be divided into positivity and negativity. If these things are necessary, their necessity is not logical, but factual, and hence arbitrary.

Law and notion are still contingently related, which means that the phenomenal world is not yet reducible to the inner, supersensible world. The attempt to correct this discrepancy leads to the subjective moment of the dialectic. In order to free its object from contingency, the understanding declares the difference between law and notion to reside in itself. Law and notion are not different in themselves, but only for consciousness, and their difference is hereby eliminated. This procedure amounts to what Hegel names 'the explanation model' (*das Erklären*), and consists in equating

¹⁶PS, 89 / HW 3, 119

the law's factual necessity with the notion's logical necessity: a law has this specific form because it is the immediate translation of its notion; and the notion is nothing other than the need for this kind of translation. Hence, for instance, the reason why electricity is divided into opposite poles lies in a notion defined precisely by such a division. In other words, electricity is divided into opposite poles because it is divided into opposite poles. Law and notion are equated, but in a tautological way, which reveals nothing concrete about either of them. The ruse of this strategy consists in positing two notions that are not really different in order to decree, afterwards, their coincidence.

Hegel is here criticizing what he views as the general one-sidedness of conventional scientific thought, based on a simple understanding of identity and difference. Since the alternative between law and notion is still construed in simple or abstract terms, consciousness can only unite the two by reducing them to a tautology. But what now becomes clear is that this union can only be brought about by overcoming the identity model favoured by the understanding. Instead of construing difference as the *external difference* between two selfsame units, artificially cancelled or 'explained' away, consciousness must look for difference *within* identity—it must come to realize that difference is really *internal difference*, or that identity is intrinsically *self-differentiating*.

This realization takes place during the dialectic's third and final stage. Initially, the equation between law and notion leads to a tautology. But the latter, on closer inspection, is more than a simple self-coincidence. While the explaining procedure, as regards the object of the understanding, 'gives rise to nothing new', as regards the understanding itself it reveals 'the very thing that was missing in the law, viz. the absolute flux itself.'¹⁷ 'What is present here is not merely bare unity in which *no difference* would be *posited*, but rather a *movement* in which a *distinction is certainly made but*, because it is no distinction, *is again cancelled*.'¹⁸ In order to demonstrate the necessity binding notion and law together, the understanding is forced to posit their difference and negate it afterwards. In doing so, however, it fails to see the necessity of these two operations. Focused on guaranteeing the unity of the inner world, it does not yet distinguish what is really necessary in the whole procedure: namely, the very movement whereby difference is *simultaneously* posited and cancelled. But the recognition of this necessity is now revealed to the understanding: 'In the process . . . of explaining, the to and fro of change, which before was outside the inner world and present only in the appearance, has penetrated into the supersensible world itself. . . . The Understanding thus *learns* that it is a *law of appearance itself*, that differences arise which are no differences, or that what is *selfsame repels* itself from itself.'¹⁹

¹⁷PS, 95 / HW 3, 126

¹⁸Ibid.

¹⁹PS, 95f. / HW 3, 126f.

The strategies devised by consciousness to deal with the collapse of thinghood—the notion of force, the difference between force's expressed and unexpressed form, the opposition between a soliciting and a solicited force, the contrast between the world of appearance and an inner world, the antithesis between universal and particular laws, and the difference between law and notion—were all grounded in a dualism that kept identity and difference on opposite sides. To be sure, inasmuch as all of these solutions were born out of the sublation of thinghood, they were no longer based on the rigid atomism of the first and second chapters. Nevertheless, this reciprocity was still anchored in two different poles, perceived as the primary source of movement. Now, on the contrary, movement itself emerges as the primary element. It is no longer an effect, or a result, but the *universal law* whereby anything whatsoever comes to be known.

While in all the previous stages difference was construed as the disagreement between selfsame units—sensations, things, forces and laws—such units are no longer admissible. Wherever one looks, difference has taken over. It is the fabric in which reality is sewn. This discovery provides an answer to the old debate concerning the origin of difference, viz. the issue of how difference can be said to emerge or issue forth from identity. For Hegel, this question is already biased: difference is not external to, but embedded in, the notion of identity. The Parmenidean $\epsilon\nu$ is therefore a mere abstraction, a divided moment whose otherness has been temporarily negated. Like all independent unities, it can only be thought of as such by being artificially removed from the infinite flow of identity and difference.

With the universalization of difference—with *infinite* difference, as Hegel also calls it—the understanding's journey comes to an end. 'Law completes itself into an immanent necessity, and all the moments of [the world of] appearance are taken up into the inner world.'²⁰ Whereas initially the sensible world was regarded as the real world and the supersensible world as an empty and unknowable beyond, what now comes to light is the opposite. The sensible world is unmasked as the mere reflection of a law issued by consciousness, a law that posits the universal difference of natural phenomena. Accordingly, the phenomenal realm no longer exists in its own right, as something contingent or arbitrary, while the understanding struggles to translate it through ideal rules and patterns. The emptiness that was first placed on the 'inner' side of things is now transferred to the outside world and eliminated.

At this stage, all that is left for consciousness is to return to its initial syllogism and draw out the final consequences. Since the phenomenal world (R') has lost its independence and vanished into the inner world (R), the understanding (C) and the inner world are now directly facing each other ($C = R$). Consciousness realizes at last that the absolute interlacement of identity and difference is not an attribute or a feature of the objective world, but the founding truth of objectivity. It realizes, in other words, that its true object is nothing other than its own law, or rather *consciousness itself*. Therefore, 'behind the so-called curtain which is supposed to conceal the inner world, there is nothing to be seen unless *we* go behind it ourselves,

²⁰PS, 99 / HW 3, 131

as much in order that we may see, as that there may be something behind there which can be seen'.²¹ The dichotomy is finally lifted and a new shape of consciousness is born. The understanding now knows that its object, i.e. reality itself, is really only its own self: *the truth of consciousness is self-consciousness*.

7.2 Selfhood as Recognition

In the previous phenomenological stages, reality was placed outside consciousness and translated into a series of increasingly sophisticated objects, corresponding to a series of increasingly sophisticated worldviews. Gradually, though, what was deemed true in itself turned out to be so only for consciousness and the realm of objectivity passed over into subjectivity. Consciousness found out that the truth of its previous objects is its own self and that all knowledge is really self-knowledge.

At this point, the whole progression might come to an end. If the Kantian chasm between cognition and reality has indeed been surmounted, and if consciousness is now simply the coincidence of the self with itself, the truth searched for in the previous stages seems to have been found. The simple empiricism endorsed at the outset resolved into a radical form of idealism and what lies ahead must be some form of self-contemplation.

This is not the case, however, because objectivity was not simply eliminated, but sublated. The objective world was negated, but also preserved as a moment of self-consciousness and raised to a new cognitive stage. To grasp the meaning of this transition, let us return briefly to the understanding's final metamorphosis. When difference was raised to absolute difference, the inner world was revealed as the universal law of the understanding, i.e. as consciousness itself. Yet precisely because this law is the law of absolute difference, the coincidence between consciousness and reality is not a simple or immediate coincidence. Consciousness is reunited with reality, but this reunion takes the form of a dynamic, or *infinite* exchange: on the one hand, consciousness distinguishes itself from reality and posits the latter as its object; on the other hand, it is immediately aware that this object is nothing other than itself, and cancels the initial distinction.²² In this toing and froing, self and other are both equally essential and completely interdependent. As if facing a mirror, consciousness is only conscious of itself by putting an 'other' ahead of itself and distinguishing itself from it; but this positing is only possible because consciousness is already conscious of itself to begin with—which implies, in turn, the sublation of a previously posited 'other', and so on indefinitely.

²¹PS, 103 / HW 3, 135f.

²²The movement through which consciousness distinguishes itself from itself and the counter-movement through which it returns to itself are not chronologically successive. At any given moment, consciousness is already simultaneously in the process of dividing itself and reuniting with itself. It is this absolute reciprocity, reminiscent of Fichte's self-identity *Trieb*, that Hegel wants to convey when he speaks of the *immediate* nature of this double movement.

Unlike a thing or an object, self-consciousness is both identical to, and different from, itself. It is what it is through the suppression of what it is not. As Hegel puts it, the 'I' of which identity is predicated is nothing but the 'return from *otherness*'²³, or *what is left* when all otherness has been suppressed; but since this 'I' is defined negatively by what it is not, it is doomed to return to its opposite and distinguish itself from it. In this circular movement, the sensuous world is preserved for consciousness, but no longer as a positive or independent world. It is 'an enduring existence which, however, is only *appearance*, or a difference which, *in itself*, is no difference.'²⁴ We have here the same syllogism discussed in the previous stage, save for one important difference. Whereas previously the syllogism had for its middle term appearance and for its extremes consciousness and reality, here the middle term is also appearance, but both extremes are taken over by consciousness. In other words, consciousness is now the unity of the self with itself mediated by appearance. In order to find itself, it must return to the sensuous world, expose appearance *as* appearance and thereby eliminate it. This movement, at this stage of the progression, is called 'desire' (*Begierde*).

The notion of desire is the starting point of a new dialectical cycle—the dialectic of selfhood, in which the 'I' will be raised from a mere vanishing point to a more substantial form of identity. This new set of developments will show that what consciousness knows in knowing itself is really much more than it initially supposes. Instead of a simple, self-referential being, lost in the pure contemplation of itself, self-consciousness is already a movement, a *Trieb*, or a peculiar form of desire. The truth of selfhood lies therefore in the examination of this movement, which requires both the analysis of different definitions of desire and the analysis of the different objects required to satisfy them. As usual, Hegel's exposition follows two simultaneous threads, focusing on the subjective and the objective implications of each phenomenological transition.

Initially, self-consciousness is '*desire in general*'.²⁵ It is the *desire to be itself*, or to raise itself out of otherness. On the subjective side, this movement is accomplished by suppressing objectivity and re-establishing the unity of the self with itself. However, this procedure is inherently self-defeating: the satisfaction of desire can only be obtained by reviving, again and again, the original source of desire. On the objective side, the same circularity is repeated, but in the opposite direction: if otherness is still nothing more than the opposite of self-identity, the object is merely a 'not-I', bound to be eliminated.

Unable to find satisfaction in this way, consciousness moves on to a new dialectical equation. This time, the transformation is brought about through the relationship's objective side. For desire to be satisfied, objectivity must be construed as something more than a mere abstraction, or a fleeting moment. 'Self-consciousness which is simply for itself, and directly characterizes its object as a negative

²³PS, 105 / HW 3, 138

²⁴PS, 105 / HW 3, 139

²⁵Ibid.

element, . . . will therefore, on the contrary, learn through experience that the object is independent.' It will learn that the object has 'not merely the character of sense-certainty and perception, but [is] being that is reflected into itself', or 'a living thing'.²⁶

The dialectical contrast between life and death stands, in Hegel's writings, for the wider conceptual contrast between an abstract or tautological mode of identity and a dynamic or dialectical one. In the element of organic life, identity and difference are not rigid, but fluid determinations. They are the limiting edges of the ceaseless alternation between union and dissolution, concentration and dispersion, growth and decline. Accordingly, living organisms are not mere things, or collections of things, but self-moving and self-determining beings. Each organism is *at the same time* the unity of its members and their self-differentiation, the universal fluidity of life and the division into new members and particular beings.

It should be noted, however, that Hegel's description does not apply simply to a new class of objects. What is here at stake is not a change of focus from a universe of 'dead' beings (sticks, stones, etc.) to a universe of 'living' ones (plants, animals, etc.). The term 'life' applies not to a part of objectivity, but to the whole of it, and the present transformation is therefore a global transformation. The initial conception of objectivity has been abandoned because *all* objectivity is, after all, life. Consciousness realizes that the distinction between identity and difference it had ascribed to the natural world is groundless, and that objectivity is really a dynamic, self-developing whole.

Subjectively, consciousness has also acquired a new depth. Its identity is no longer simple or abstract, but concrete or 'self-reflected'. On the one hand, the subject loses itself in the multiplicity of the objective world and is led to assume different shapes and functions. This expansion is mirrored by nature itself, which also displays a succession of independent shapes and operations. On the other hand, consciousness knows that this independence and differentiation are nothing *in themselves*. The truth of the natural world is still self-consciousness, and the independence of its different shapes a spurious independence, to be negated and reabsorbed.

Hegel's somewhat cryptical description may be illustrated by a variety of physiological examples: the need for oxygen and water felt by every living being and overcome through breathing and drinking; or the phenomenon of hunger, overcome by the consumption of plants and animals. In these cases, desire is not optional or arbitrary. The self's survival is actually dependent on the consumption and self-integration of external objects (viz. air, water, plants, animals). Yet this concession to objectivity is only temporary: consciousness is still the essential moment of the relationship, and otherness merely the means whereby the self reaffirms its sovereignty.

²⁶PS, 106 / HW 3, 139

Hegel returns to this theme in the *Encyclopaedia*, where consciousness' first contact with the natural world is also mediated by *Begierde*, and where the suppression of this initial longing depends on our ability as human beings to 'use Nature for our own advantage, to wear her out, to wear her down, in short, to annihilate her.'²⁷ In a way, the history of human science and technology is a history of desire and satisfaction. But although consciousness develops increasingly complex ways of consuming and absorbing the natural world, its ultimate end is not the object of desire, but the satisfaction it affords. The truth of the object as object of desire does not reside in itself, but in the use that consciousness makes of it. Each natural thing and each living being exist on their own right, but only because they exist *for consciousness*: water, before it is a chemical compound, is 'that which allows human beings to slake their thirst'; plants and animals, before they are independent living organisms, are 'that which allows human beings to quench their hunger.' Likewise, the end of desire is 'our satisfaction, or our self-feeling, which had been disturbed by a lack of some kind or another. The negation of myself which I suffer within me in hunger is at the same time present as an other than myself, as something to be consumed; my act is to annul this contradiction by making this other identical with myself, or by restoring my self-unity through sacrificing the thing.'²⁸

In the case of hunger, this 'return to selfhood' is achieved in a literal way: animals and plants, in being consumed and digested, become 'flesh of my flesh'. Their identity is merged in my own identity and becomes the subject of a new desire, aimed at a new object. But this movement, repeated over and over again, rests on a circularity that is similar to the one found in the previous model. The end of desire is indeed the satisfaction obtained through the suppression of objectivity, but for this suppression to take place there must be a world to suppress. Just like before, the very form of the relationship leads to the continual resurrection of the original source of desire. Self-consciousness is doomed, therefore, to an eternal state of dissatisfaction—a fate which Hans-Georg Gadamer, in his reading of this chapter, compared to the perpetual oscillation highlighted in Goethe's *Faust*:

Thus I reel from desire to satisfaction,
And in satisfaction I long for desire.²⁹

To overcome this vicious circle, self-consciousness must abandon this model and redefine the notion of desire. And this transformation comes about, once again, through the objective side of the relationship. The satisfaction of thirst and hunger are necessary conditions for the maintenance of self-consciousness, but they are not yet its essence or truth. If the self's return to itself is not to end in circularity, it must be countered by an equally strong symmetrical movement, capable of fighting back, as it were, and opposing subjectification. This kind of resistance cannot be offered by plants or animals, for the latter are ultimately but mere means of satisfaction. This

²⁷ PN, 5 / HW 9, 13

²⁸ PN, 5 / HW 9, 14

²⁹ Goethe, *Faust. Erster Teil*, ll. 3249f. See Gadamer, *Hegels Dialektik*, 53f.

counter-movement can only be ensured by an object that is no longer a means, but an end in itself—an object that ‘is in its own self negation, and in being so is at the same time independent.’³⁰ In other words, this counter-movement can only be ensured by *another subject*, or by *another self-consciousness*.

In this crucial turning point, the subjective and the objective moments are finally brought together. The ‘I’ is no longer faced with a ‘not-I’, destined to be consumed and subjectified. Its object is now ‘another I’, which is itself faced with a self-conscious subject. This development re-enacts, at a higher level, the self-duplication of force, examined in the previous section. Just as force’s expression was at first solicited by an external element, and found afterwards to depend on a second symmetrical force, the self now breaks up into two selves, whose identity is irrevocably intertwined. Each subject, in being conscious of itself, is simultaneously conscious of the other subject’s self-consciousness.

This transformation is encapsulated in two short sentences, which mark the beginning of a new dialectical cycle:

*A self-consciousness exists for a self-consciousness. Only so is it in fact self-consciousness; for only in this way does the unity of itself in its otherness become explicit for it.*³¹

But this is not all. This first revelation is immediately followed by a second one, whose full implications will only become apparent much later. Just as the play of forces was not reduced to two forces, but extended to a universal web of forces, so too will the dialectic of self-consciousness eventually lead to a similar universalization. As the remaining stages of the progression will render more explicit, to speak of *two* mutually related selves is already to speak of a *global system* of mutually related selves. The discovery that singularity is rooted in universality, or that subjectivity is rooted in intersubjectivity, is the first defining feature of the Hegelian notion of Spirit, or *Geist*:

A self-consciousness, in being an object, is just as much ‘I’ as ‘object’. And with this, the notion of *Spirit* is already available to us. What still lies ahead for consciousness is the experience of what Spirit is—this absolute substance which is the unity of the different independent self-consciousnesses which, in their opposition, enjoy perfect freedom and independence; ‘I’ that is ‘We’ and ‘We’ that is ‘I’.³²

This new thesis, one of Hegel’s most original and controversial insights, represents a complete reversal of our usual conception of personal identity. Instead of an original, independent self confronted with the presence of an other, Hegel argues that this other is the condition for the self to emerge in the first place. Intersubjectivity is not a result, arising from the mutual interaction of simple autonomous selves, but the founding instance of personal identity. Self-consciousness is an inherently social phenomenon.

³⁰PS, 109f. / HW 3, 144

³¹PS, 110 / HW 3, 144f.

³²PS, 110 / HW 3, 145

Yet this idea is anything but self-evident. Indeed, it is directly opposed to the very experience of subjectivity: if my own self is the innermost core of my existence, something only I am truly aware of, how can it depend on a social encounter? Furthermore, if there is no self prior to otherness, how can the acknowledgement of the existence of an ‘other’ ever come about? These and other perplexities bear witness to the persistence of the *Phenomenology*’s previous dichotomic framework. Any interpretation of this passage that assumes the priority of selfhood over sociality entails a return to the very misconception that is now being criticized. And yet this is precisely the kind of approach our standpoint tends to fall back on. The claim that consciousness is only self-conscious in the wake of a social exchange is likely to be read in light of the simple contrast between realism and solipsism. Since our usual account of reality is dependent on the belief in the actual existence of other people, on the assumption of a shared ground of natural qualities, on a collectively agreed system of linguistic conventions, etc., intersubjectivity is certainly an essential feature of personal identity. But this kind of reading still allows for the possibility of a primary core of selfhood, immediately intuited and ultimately independent from the social sphere.

According to this version of the argument, Hegel would be merely pointing out, like Aristotle, that each of us is a ‘political animal’ (ζῷον πολιτικόν); that our identity is just as much defined by our sense of selfhood as it is by our social context, circumstances and encounters. Now all of this may well be true, but Hegel’s point is much more radical. Sociality is no longer to be envisaged as something merely added to self-consciousness, as one of its predicates or modes of expression. Rather, it entails the very dissolution of the antithesis between selfhood and self-expression. Accordingly, self-consciousness ‘is *in* and *for itself* when, and by the fact that, it is in and for itself *for another*; that is, it is [what it is] only in *being recognized as such* [als ein Anerkanntes].’³³

This mutual act of recognition, whereby each subject becomes what it is, is the basis for all of the *Phenomenology*’s further developments. Henceforth, the dialectical grammar of self-identity, with its different and increasingly complex declensions, will coincide with the self-evolving grammar of human relations. The same demand for recognition will give rise to a long succession of cognitive shapes—from the original encounter between two recognizing selves to more complex forms of social interaction, and to the shifting roles of master, slave, citizen, vassal, believer, actor, artist and philosopher. For Hegel, ‘it is the nature of humanity to press onward to agreement with others; human nature only really exists in an achieved community of consciousnesses’³⁴, and the phenomenological progression is the long road towards that achievement.

Before moving on, however, let us return to our initial query and determine how this new transformation fits into the more general debate concerning the relationship between freedom, power and knowledge. At the beginning of the *Phenomenology*,

³³PS, 111 / HW 3, 145

³⁴PS, 43 / HW 3, 65

consciousness was led to acknowledge the contradictions embedded in its usual mode of cognition and driven to a 'state of despair.' As was previously suggested, its original condition, 'hampered' by 'the so-called natural ideas, thoughts and opinions'³⁵, can be compared to that of the prisoners held inside Plato's cave, and its initial self-criticism to the breaking of their chains. Afterwards, throughout the first three chapters, consciousness begins its ascent. The dialectic of objectivity is fuelled by the disagreement between reality *in itself* and reality *for consciousness*, or truth and appearance: whenever consciousness believes it has found the truth about its object, the latter turns out to be something different, and what seemed like the end of the path is revealed as a new beginning. This experience, re-enacted at different cognitive levels, leads to the emergence of self-consciousness, that is, to the complete integration of objectivity within the sphere of subjectivity. This turning point offers the first, albeit provisional solution to Plato's predicament. By abolishing the difference between reality and appearance, or between an 'inner' and an 'outer' world, consciousness enters 'the native realm of truth'.³⁶ The *doxastic* assumptions of the previous stages seem to have been left behind, and a new, *epistemic* chapter is about to begin.

The experience of self-consciousness then shows that this victory is only apparent, since the 'I' is still irrevocably bound up in objectivity, still stranded to the circular logic of desire. What follows, however, is not simply the collapse of another mode of cognition, but the introduction of a new philosophical theme, which seems to distance Hegel from Plato and the critical tradition. Indeed, the dialectic of desire leads to what Axel Honneth dubbed an 'intersubjectivist innovation'³⁷, which transcends the idealist framework previously set up by Hegel and paves the way for what is arguably the *Phenomenology's* most original philosophical contribution.³⁸ According to Alexandre Kojève, one of the first Hegelian exegetes to acknowledge the revolutionary nature of this transition, 'in having discovered the notion of recognition, Hegel found himself in possession of the key idea of his whole philosophy'.³⁹

³⁵ Ibid.

³⁶ PS, 104 / HW 3, 138

³⁷ Honneth, *Kampf um Anerkennung*, 31

³⁸ Hegel's argument is not altogether without precedent. The idea that self-consciousness is necessarily mediated by a social exchange is already found in Fichte, both in the *Wissenschaftslehre* and in the *Grundlage des Naturrechts*, where individuality is defined as a 'reciprocal' or 'communal' concept, which 'can only be conceived in relation to another', and where personal freedom is likewise grounded in a reciprocal or communal interaction. (Fichte, *Sämmtliche Werke*, vol. 3, 47f.)

³⁹ Kojève, *Hegel, Marx et le christianisme*, 352. This 'key idea' is crucial to the classical debate about the continuity of Hegel's text. The advent of self-consciousness signals not only an inversion of the conceptual relationship between subjectivity and objectivity, but also a transformation of the *Phenomenology's* thematic scope: instead of the dialectical evolution of epistemologically-oriented 'shapes of consciousness', Hegel now focuses on concrete self-conscious individuals, their mutual interaction and their physical and social surroundings.

Whether or not one endorses Kojève's sweeping claim, the conceptual change introduced by the dialectic of recognition is hardly innocuous. Whereas the *Phenomenology* was heretofore fuelled by the tension between subjectivity and objectivity, the focus is now placed on the tension between subjectivity and intersubjectivity. But this new process does not imply the elimination of the previous one. In the next chapters, the progression will deal simultaneously with two main themes: as we will see below, the transition from the stage of self-consciousness to the stages of Reason and Spirit will lead to the acknowledgement both that the subject is not independent from the object and that the subject is not independent from other subjects. Spirit is consciousness that is 'conscious of itself as its own world, and of the world as itself'⁴⁰; but it is also 'the unity of the different independent self-consciousnesses which, in their opposition enjoy perfect freedom and independence.'⁴¹

7.3 Masters, Slaves and Philosophers

When the notion of recognition comes on the scene, it is not fully understood. As always, the progression is slow and surgical, so as to exhaust the argument's various implications. Initially, each self-consciousness sees itself as the essential moment of the relationship—as a self that *recognizes* faced with a self that is *recognized*. Still guided by the unilateral logic of desire, each regards the other as a mere object, part of the natural world it seeks to negate. But since the truth of self-consciousness is now a purely reciprocal exchange, this asymmetry is no longer warranted. To become truly self-conscious, each subject must appear to the other as the opposite of a natural, inert object. Each must be recognized as a pure act of recognition, which can no longer be seized or objectified. For this to happen, however, each must undertake 'the pure negation of its objective mode' and show 'that it is not attached to any specific *existence*, . . . that it is not attached to life.'

Hegel's argument hinges on the distinction between a natural life, determined by a natural drive for self-preservation, and an autonomous or self-determining life. In the first case, the individual's attachment to concrete existence is similar to that of animals and plants, and he is therefore bound to be recognized in similar terms, that is, as a vehicle for satisfaction and self-preservation. In the second case, self-preservation gives way to self-determination and to a new conception of self-identity. Each individual shows the other he is willing to risk his own life in order to be recognized.⁴² And since this demonstration occurs simultaneously on both sides of

⁴⁰PS, 263 / HW 3, 324

⁴¹PS, 110 / HW 3, 145

⁴²For practical reasons, I follow Hegel's use of masculine pronouns throughout this chapter. This does not mean that the master-slave dialectic applies exclusively to men. As already pointed out, it refers to a cognitive model adopted by human consciousness as such.

the relationship, what takes place is a struggle for recognition where both parties seek the death of the other and the confirmation of their own independence.⁴³

This mortal struggle is the most basic form of social interaction. It is the first means devised by self-consciousness to attain recognition. Yet if either contender succeeds in eliminating the other, recognition itself comes to an end. Although the winner will have risked death, as required, he will no longer be able to recognize, or be recognized by, the other, and his own selfhood will sink back to immediacy. Such victory will be an empty victory, leading back to the previous phenomenological standstill.

The dialectical link between both contenders cannot be severed because self-identity and otherness are equally essential to the relationship. This initial strategy must therefore be revised. Although each self-consciousness must assert its independence by negating the independence of the other, each can only become independent by being recognized as such. Consequently, each must attempt to reduce the other to a state of nothingness that nonetheless preserves its ability for recognition. In other words, the struggle must not end in death, but in domination: the vanquished party must be allowed to live and forced to submit to the winner and recognize his independence.

The contrast between life and death thus gives way to the contrast between sovereignty and servitude. The struggle unto death is replaced by the confrontation between a master, or an 'independent consciousness whose essential nature is to be for itself',⁴⁴ and a slave, or a 'dependent consciousness whose essential nature is simply to live or to be for another.'⁴⁵ Initially, what binds the slave to the master is the fear of death. Both have risked their lives in the struggle, but their willingness to sacrifice themselves was unequal. Whereas the master, in the face of death, refused to give up his independence, the slave chose to submit and save himself. His attachment to life is what holds him in bondage, the 'chain from which he could not break free in the struggle.'⁴⁶

Self-consciousness is now construed as a synthesis between dependence and independence, divided between a sovereign and a slavish standpoint. Accordingly, the initial contrast between subjectivity and objectivity is no longer mutually

⁴³ It is tempting to interpret this mortal struggle as a depiction of humankind's remotest infancy, especially given the close similarities with Hobbes' account of a primordial 'state of nature', famously defined as a 'war of all against all'. At stake, however, is not a historical or anthropological description, but a specific cognitive model. In the stage of recognition, consciousness adopts a new perspective on its own understanding of reality, and more precisely on what the most immediate form of knowledge amounts to. It maintains that the most primitive form of cognition is not sense-certainty, nor perception, nor the understanding, nor self-consciousness as such, but a struggle between two mutually recognizing selves. And though consciousness is not actually reduced to this primitive struggle, all of its thoughts and judgements are now perceived as immanent developments thereof.

⁴⁴ PS, 115 / HW 3, 150

⁴⁵ Ibid.

⁴⁶ PS, 115 / HW 3, 151

exclusive and the natural world is readmitted into the relationship. The model becomes thereby a triangular model: apart from fear, the relationship between master and slave is also mediated by 'thinghood', that is, by the objective element which the master managed to overcome and which the slave remained attached to during the struggle. This 'thing', the material world, is now the point of contact between both extremes, viewed differently from each side.

The master engages in a mediate exchange with the other edges of the triangle: his relationship with the slave is mediated by the thing, inasmuch as he is now the thing's proprietor and forces the slave to work on it; likewise, his relationship with the thing is mediated by the slave, inasmuch as the latter transforms the thing and makes it ready to be consumed and enjoyed by the master. On the slave's side, the relation is also twofold: the slave can only secure the master's mercy (and thus his own survival) through his work on the thing; however, his access to it is granted by the master, who owns both the thing and the fruits of the master's work.

Hegel undertakes a systematic survey of the model's internal contradictions, in order to reveal its inherent instability. As regards the master, although he succeeded in subduing his rival and obtaining recognition, his victory is not as complete as he might have expected. To be sure, he managed to rise above objectivity and to achieve the satisfaction denied in the dialectic of desire. The slave's work, demanded by the master in exchange for his survival, transformed the master's existence into one of consumption and enjoyment. However, his independence is primarily grounded in the recognition afforded by the slave, of which consumption and enjoyment are the indirect results. Recall that the main condition for the achievement of independence, at the beginning of the dialectic, was the recognition of each self on the part of another recognizing self, equally eager to prove his independence. Only two rivals willing to put their freedom above their particular existence were able to grant each other the independence they both craved. This boldness, however, is precisely what the master no longer finds in the slave. After the struggle, the slave relinquished his independence and chose a life of subjection. Therefore, the recognition he is able to offer is no longer the recognition of a free, independent self-consciousness.

From the point of view of the slave, the situation is no less ambiguous. Although his subjection seems at first absolute, his defeat is the result of a thwarted claim to independence whose essence is very similar to the master's. After the struggle, the slave was consigned to a position of inferiority, but he also staked his life and also faced the possibility of death. He was fearful 'not of this or that particular thing, or just at odd moments, but [his] whole being has been seized with dread'. In this experience, he 'has been quite unmanned, has trembled in every fibre of [his] being, and everything solid has been shaken to its foundations'⁴⁷. Now it should be recalled, once again, that this radical self-detachment, this 'rooting-out of all immediate being', is the very movement out of which self-consciousness is born. To rise

⁴⁷PS, 117 / HW 3, 147

above particularity, self-consciousness must abandon the realm of natural existence and divest itself of every trace of contingency. Only as absolute negativity, as 'pure being-for-self', can it find its truth.

By coming face to face with death, the slave is left, but for one split second, with nothing more to lose. But this foretaste of freedom is soon cut short. His defeat forces him to submit once more to the contingency of natural life. Yet whereas the master's standpoint leads to a dead end, the slave's standpoint is more promising. He is now free *in posse*, and his work will eventually allow him to break his chains and embrace a new mode of identity. Servile labour will bring about the synthesis, within the same self-consciousness, of the master's and the slave's opposite stances.

Here, two brief remarks are in order. First, to grasp the meaning of this new transition it is useful to bear in mind Hegel's preliminary claims about the progression's immanent necessity. As always, he is not simply hazarding a set of possible scenarios, which may or may not ensue from the situation at hand. He is spelling out what he considers to be the necessary outcome of the previous stage's inner contradictions. And although it is certainly possible that servile labour remains servile, and that the master's domination is perpetually maintained, this possibility should be put down, following Hegel's logic, to self-consciousness' inability to grasp the implications of its own nature. The slave's labour harbours a dialectical tension whose logical outcome is freedom and independence. And, likewise, the master's sovereignty harbours a dialectical tension whose logical outcome is servitude. But whether or not the latter takes place is a matter for speculation. The *Phenomenology* follows the slave's self-development, but says nothing about the master's ultimate fate.⁴⁸ It is up to the slave, rather than the master, to carry the phenomenological torch.

Second, since the master's authority stems from the slave's activity—since his enjoyment is produced and determined by the slave—one might expect the master-slave dialectic to culminate in a simple role reversal: the slave would end up recognizing his true power and rising to sovereignty, while the master would succumb to his own impotence.⁴⁹ But although this transformation is undoubtedly possible—Marx and Engels would later argue that it is the motor of human history—it is still but an immanent development of the same phenomenological model. It may be re-enacted again and again without ever questioning the basic premises upon which the model is built: new masters will be subjected to new slaves, the latter will give way to new masters, and so on indefinitely. Hegel avoids this impasse by reversing this basic strategy: the sublation of the master-slave scenario is not achieved through the slave's rebellion, but through his submission. By abandoning his initial demand for recognition and exploring his new condition, the slave will undergo a more profound metamorphosis and discover a more sophisticated kind of freedom.

⁴⁸ Although Hegel suggests in the *Encyclopaedia* that the slave's liberation entails the liberation of the master. See PM, 162 / HW 10, 227.

⁴⁹ See Gadamer, *Hegels Dialektik*, 59f.

At first, the slave's independence is only the seed of independence, planted inside him by his brief confrontation with death. Although his attachment to life proved stronger than his drive for independence, the former will now be eradicated and the latter will gain the upper hand. This process is once again twofold, divided into an objective and a subjective moment. Objectively, the concrete action of handling and reshaping external things enables the slave to oppose the independence of the natural world. Through manual labour, he draws the objects out of their immediate existence and grants them a specific form and meaning. Subjectively, the slave comes to acknowledge this transformation as his own doing and to find therein a new source of independence. His 'formative activity is simultaneously the individuality or the pure being-for-self of consciousness, which now, in the work outside of it, acquires an element of permanence; it is in this way, therefore, that the working consciousness comes to see in the independent being [of the object] its *own* independence.'⁵⁰

What Hegel has in mind is a peculiar form of transfer, through which the independence of the object is progressively recognized as the slave's own independence. More than manual labour in general, which can be carried out in an automatic and thoughtless way, this procedure is perhaps best illustrated by the work of a craftsman or an artist. Initially, the object is nothing more than pure materiality: a brute presence, hostile to any kind of transformation. The worker has to adapt to the object's mode of being, yield to the peculiarity of its nature. But as his efforts increase, the object's materiality retreats to the background and his design is brought to the fore. At the end of the process, the object is no longer a mere thing, but an artefact or a token of the worker's identity.

Like the master before him, the slave has now restored his independence vis-à-vis the natural world. But while the master's independence rested on a preformed object, worked on by the slave, the latter is the author of his newly shaped world. The difference between these two attitudes reflects, once again, the *Phenomenology's* recurrent distinction between a simple or abstract conception of negativity and a concrete or dialectical one. Whereas the master's enjoyment amounts to a simple negation of objectivity, devoid of movement or mediation, the slave brings about a determinate negation of objectivity. He alienated his own identity, but only to re-integrate it in a wider dynamic whole.

Looking back on this entire movement, it is easy to see why the *Phenomenology* played such a decisive role in the development of modern social theory. Hegel's dialectic of self-consciousness paved the way for the Marxist tradition not only because of its focus on the reversibility of power, but also because of its emphasis on labour. For Marx, as for Hegel, labour represents a force far superior to the master's sterile domination. For both thinkers, it is up to the worker to take possession of his work and overcome his enslaved condition. Unlike Marx, however, Hegel is specifically interested in the slave's *voluntary* subjection. If he is to win back his freedom, he must first give it up altogether. The object of his work must appear to him as in a mirror, endowed with the same independence he himself longed for before the

⁵⁰PS, 118 / HW 3, 154

moment of subjection. If this initial contrast is not radical enough, the transfer of power will not be complete. Work may lead to a material liberation, but not to a true emancipation.

True freedom requires, on the contrary, absolute fear. Only through a visceral, all-encompassing dread can work bring forth the transformation envisaged by Hegel:

If consciousness *forms* the thing without the initial absolute fear, it is only a vain, self-centred attitude; for its form or negativity is not negativity *in itself*; and therefore its formative activity cannot give it a consciousness of itself as essential being. If it has not experienced absolute fear but only some lesser dread, the negative essence has remained for it something external, its substance has not been infected by it through and through.⁵¹

But fear in itself is not enough. If the slave succumbs to a state of paralysing dread, his liberation is also compromised. Only servile work can turn his fear into a positive and cathartic force. Accordingly, Hegel speaks of the ‘discipline of service and obedience’⁵², in a way that has understandably put off many modern readers. Yet although his emphasis on servitude is indeed controversial, one must not lose track of the argument’s scope, nor of its significance for the *Phenomenology* as a whole. Hegel is not referring to an actual historical situation, nor advocating servitude as such. He is arguing that the notion of servitude is a necessary stage in consciousness’ immanent path of self-recognition. His argument is not political, which is why all references to Marxism, contractarianism or authoritarianism must be handled with caution. As always, to unmask each stage’s inner inconsistencies, consciousness has to experience them and acknowledge their effects. Just as a mortal struggle was the logical follow-up of desire, so too servitude follows from consciousness’ view of itself as a blind demand for recognition. In like manner, the next phenomenological stage will reveal that the present model is itself in need of correction. The freedom achieved through servitude is but a transient, self-sublating construct, whose logical outcome is a different and more sophisticated form of freedom.

It should also be noted that the kind of servitude Hegel refers to at the end of this section is no longer directly caused by the domination of the master over the slave. Even though the former emerged initially as the winning party, the latter is eventually revealed as the true protagonist of the relationship. For the slave exerts on himself, voluntarily, the action that the master had forced upon him, and by doing so drives the master into obsolescence. It is the slave who, in order to save his own life, renounces his independence and embraces a life of servitude. His ‘absolute master’, as Hegel puts it, is not the master, but life itself.⁵³ The master-slave dialectic is therefore the dialectical struggle of the slave with himself, in which the master is merely a kind of trigger: his role is to exert the pressure necessary for the slave to retreat into himself and re-emerge as an independent self-consciousness.

⁵¹ PS, 119 / HW 3, 154f. Translation modified.

⁵² PS, 119 / HW 3, 154

⁵³ PS, 117 / HW 3, 153. Translation modified.

Once the slave recognizes the world's independence as his own independence, self-consciousness is elevated to a new stage. This development seems to replicate, at first, the discovery made at the beginning of the chapter, namely that self-consciousness is the truth of consciousness, or that subjectivity is the truth of objectivity. However, while initially the 'I' was purely evanescent, or the mere 'return from otherness', it is now conscious of being the truth of that otherness. The dissolution of objectivity is no longer something that merely happens, as a result of the self's inner contradictions, but something that is actively brought about by consciousness. For the 'I' is now a *thinking* self, to whom the outer world is the content of a free, manipulable set of thoughts. Whereas in the previous stages the relationship with otherness was still one of subordination, grounded in different forms of desire and recognition, consciousness has succeeded in breaking free. Instead of yielding to objectivity, or adapting to a preformed version of things, it is now the primary source of reality's identity, the standard with which to assess its meaning and value.

By subordinating the outer world to the world of thought, the phenomenological subject rises for the first time to a *philosophical* stance. Its object is no longer objectivity as such, but its own perception of it, and its thoughts are directed, therefore, at thinking itself. 'In thinking', Hegel argues, 'I *am free*, because I am not in an other, but remain simply and solely in communion with myself'. Hence, 'my activity in grasping [something] [*in Begreifen*] is a movement within myself.'⁵⁴ This passage calls upon the distinction between 'representational thinking' (*vorstellendes Denken*) and 'notional thinking' (*begreifendes Denken*) discussed in the preface.⁵⁵ As we have seen earlier, these two cognitive attitudes correspond to two different ways of understanding the relationship between subjectivity and objectivity. In the first case, the subject is conceived as a simple, selfsame 'I', confronted with a series of simple, selfsame objects, and both sides remain the same before and after the relationship, i.e. mere abstractions. In notional thinking, the relationship is concrete or dynamic. The object is no longer perceived as a mere object, but as the subject's true nature, and the link that binds them is not contingent, but necessary.

Whereas 'what is [merely] represented [*das Vorgestellte*] . . . has, as such, the form of being something other than consciousness', and is thus set up as the essential moment of the relationship, a notion (*ein Begriff*) is 'a content grasped *in thought*', which renders consciousness '*immediately* aware of its unity with [it].'⁵⁶ In the progression's previous stages, the first of these attitudes was continually challenged, but not altogether eliminated. With this new transition, however, notional thinking achieves an important victory. Self-consciousness is no longer attached to otherness as such, or to a world that merely is. Since everything is now the correlate of a specific thought, everything that is *is so* insofar as it is *thought to be so*.

⁵⁴PS, 120 / HW 3, 156

⁵⁵See Sect. 6.2.

⁵⁶PS, 120 / HW 3, 156. Translation modified.

Stoicism is the first phenomenological incarnation of this attitude. To the stoic, reality is either true or false, good or bad, pleasant or painful—but only insofar as he or she thinks it so. These attributes do not belong to reality in a substantial or intrinsic manner. They are the results of a complex web of cognitive claims (ὁπολήψεις), whose content and value are subjectively determined.⁵⁷ Stoic freedom consists therefore in the power to review and redefine the scope and configuration of this cognitive web. If reality's 'imprint' on consciousness is not absolute, but agreed upon, what usually appears natural or self-evident is merely the correlate of a specific set of thoughts, whose content can be changed. If—to quote Marcus Aurelius—'everything is opinion, and opinion is in [one's] power'⁵⁸, the world's grip on the subject is much looser than usually supposed. And once this is acknowledged, 'what is essential is no longer the difference as a *specific thing*, or as *consciousness of a specific natural existence*, . . . but only the difference posited by *thought*, or the difference which from the very first is not distinguished from myself.'⁵⁹

If reality is grounded in thought, its meaning is not fixed or inevitable, but open to change. And this change is all the more necessary, according to stoicism, since consciousness is usually led to accept as true, or good, or relevant, what lies beyond its control. Personal freedom is compromised by a series of assumptions that are not perceived as assumptions, but as facts, and stoicism's aim is precisely to denounce this misconception and restore thought's original plasticity. Stoicism highlights the need to roll back our usual expectations about reality, to reduce life's share of contingency and to limit our existential horizon to what is really up to us (ἐφ' ἡμῖν). This transformation entails a radical revision of the meaning and importance usually ascribed to external goods, physical pleasure, material wealth and all the things whose attainment cannot be guaranteed, and the pursuit of which corresponds therefore to a form of imprisonment.

The freedom professed by stoicism can only be found in the element of thought, away from the commotion of human passions and mundane achievements. This negative ideal amounts to a radical form of indifference (ἀπάθεια), which transcends both the master's material comfort and the slave's physical hardship:

Whether in the throne or in chains, in the utter dependence of its individual existence, [the stoic's] aim is to be free, and to maintain that lifeless indifference which steadfastly withdraws from the bustle of existence, alike from being active as passive, *into the simple essentiality of thought*.⁶⁰

Apart from Hegel's explicit association between this new standpoint and historical stoicism, the argumentative path followed throughout this chapter also evokes the one encountered earlier in the *Gorgias*. In Plato's dialogue, the dispute between Socrates and Callicles hinged on the dialectical interplay between freedom, knowledge and ignorance. More precisely, it revolved around the issue of whether human freedom requires a special kind of knowledge. Callicles embodied a domineering

⁵⁷ See, for example, Marcus Aurelius, *Meditations* II, 15; IV, 3; VI, 52; IX, 32; XII, 8, 22, 26.

⁵⁸ Marcus Aurelius, *Meditations* XII, 22: "Ὅτι πάντα ὑπόληψις καὶ αὐτὴ ἐπὶ σοί."

⁵⁹ PS, 121 / HW 3, 157. Translation modified.

⁶⁰ Ibid. Hegel is alluding to Marcus Aurelius, who sat 'in the throne' of the Roman empire, and to Epictetus, who was born 'in chains' and released from slavery.

attitude that is in many ways similar to that of Hegel's master: his idea of freedom was linked to pure gratification, secured through the domination of a weaker or 'baser' sort of individuals by a stronger or 'nobler' sort. Socrates' attitude, on the other hand, anticipates the servile standpoint portrayed in the *Phenomenology*. For him, inasmuch as the alleged benefits of physical strength, wealth or power praised by Callicles rest on immediate assumptions, whose truthfulness has not been proven, they are neither empowering nor liberating. They amount to mere δόξαι, or to misplaced ὑπολήψεις, which leave human beings at the mercy of what they cannot control. Because of this, Socrates rejects Callicles' idea of power and the definition of freedom in which it is grounded. Like Hegel's slave, he aspires to a different kind of deliverance, no longer dependent on concrete pleasures or concerns. His aim, like that of the Stoics, is not to dominate others, but to achieve self-mastery (ἐγκράτεια ἑαυτοῦ), or the 'simple essentiality of thought.'⁶¹

In the *Phenomenology*, this new standpoint is the logical outcome of the master-slave dialectic. Unlike the master and the slave, the stoic is aware that enjoyment and domination are not essential determinations, but only thought to be such. Yet although this discovery represents a decisive turning point in the progression, it is not yet the truth of self-consciousness. The freedom unveiled by the slave's intellectual emancipation is at first but an immediate result, or the freedom of thinking *in general*. Self-consciousness has risen above the realm of contingency and proclaimed its independence vis-à-vis the objective world, but its standpoint is now reduced to this very proclamation.

When the stoic comes across a determinate content, he immediately points out its lack of essentiality. Whatever is deemed true or good is promptly unmasked as the reflection of a subjective judgement and led back to the element of thought. The stoic refuses to abide by any particular judgement because his freedom is as wide as thought itself. It cannot be reduced to this or that specific truth without relapsing into the dogmatism of the previous stage. But the continual dissolution of particularity is only a negative movement, in which nothing positive gets determined. Although stoicism knows what the true and the good are *not*, it is incapable of determining what they are. Its ideal is thought itself, that is, a thought whose only content is the fact that it has none.

The stoic's position, as Hegel conceives it, is untenable. Thought *as such*, devoid of particularity, is also deprived of practical or existential depth. It is constitutively unable to provide a criterion with which to distinguish what is true from what is false, what is right from what is wrong, what should be pursued and avoided. And since such criteria are nonetheless needed, the stoic must find them where it cannot look for them, namely within objectivity. Since 'the notion as *abstraction* cuts itself off from the multiplicity of things, and thus has *no content in its own self*', one has to be 'given to it.'⁶² This predicament echoes the one faced earlier by the subject of desire: though stoicism rejects the essentiality of the objective world, it is forced to return to it again and again.

⁶¹ On the relationship between the *Gorgias* and Stoicism, see Alesse, 'Alcuni esempi della relazione tra l'etica stoica e Platone'; Long, *Epictetus*, 70ff.; Striker, 'Plato's Socrates and the Stoics'.

⁶² PS, 122 / HW 3, 158

The next phenomenological attitude, scepticism, calls into question not only the determinate contents that make up the objective world but also the very relationship between cognition and objectivity. Whereas the stoic preserves reality's referential status but finds its truth in the realm of thought, scepticism no longer allows for this dualism. True freedom requires a truly self-referential mode of cognition, no longer fuelled by the negation of otherness. For the first time, the objective world to which stoicism is forced to go back to is taken over by thought and nullified.

Hegel's focus on scepticism is not confined, however, to a specific phenomenological stage. The aim of the entire progression, as we have seen, is to overcome the dogmatism of natural cognition through a 'scepticism directed against the whole range of phenomenal consciousness.'⁶³ Generally speaking, every phenomenological stage leads to scepticism: each cognitive model set forth by consciousness ends up being refuted and replaced by a new admission of ignorance. But this ignorance is never final or irrevocable. It does not lead to a pure nothingness, but to a determinate nothingness, which is the starting point of a new cognitive experience. On the one hand, the entire *Phenomenology* is grounded in the conviction, inherited from Plato, that the attainment of truth can only be envisaged once all cognitive presuppositions have been identified and suspended. Philosophy must bring about 'a state of despair about all the so-called natural ideas, thoughts, and opinions.'⁶⁴ On the other hand, if this procedure is to yield a positive result, it cannot be reduced to a simple form of scepticism. The attainment of truth requires a *self-evolving* scepticism, whose results are both negative and positive.

The 'exposition of the untrue consciousness in its untruth is not a merely negative procedure.' To think otherwise is to give in to the 'scepticism which only ever sees pure nothingness', and 'cannot get any further from there.'⁶⁵ In the introduction, Hegel pointed out that this one-sided attitude is not the *Phenomenology*'s final destination, but only one of the many stops lying ahead. He was referring to the stage of scepticism and to the radical stance we are now considering. Whereas the stoic retreated into the element of thought but failed to come up with a concrete criterion of truth, the sceptic denies the existence of such a criterion and decrees the impossibility of actual knowledge. While the stoic is still attached to the possibility of a rational redemption of reality, the sceptic regards this hope as naïve: ultimately, each thought is just as good as the next one, for none is able to convey anything certain or definite about reality.

The seriousness of this charge is also reflected in the *Phenomenology*'s global architecture. Until now, the refutation of each model left behind a residue of meaning, in need of a new form of expression, and this need prompted consciousness to review its different stances and accommodate new elements and ideas. But the emergence of scepticism threatens to put an end to this pattern: in its most radical form, scepticism represents, if not the exhaustion of the phenomenological

⁶³PS, 50 / HW 3, 73

⁶⁴Ibid.

⁶⁵PS, 50f. / HW 3, 73f.

mechanism, at least the exhaustion of the method that has hitherto been followed. For the sceptical outlook entails a voluntary sabotage of the very movement of adhering to a cognitive model, whichever it may be. Instead of looking for new ways of correcting the shortcomings of a given worldview, scepticism turns its back on the whole undertaking and refuses to assent to any truth other than that there is none, or that all standpoints are ultimately similar, and equally vain.

In the first chapter, consciousness assumed the existence of a reality independent from thought and stated naïvely that things *are what they are*. Later, with the emergence of stoicism, it abandoned this naïve realism and came to acknowledge that things *are not what they are*, but only *thought of* as such. To be sure, these two standpoints are radically different, but the sceptic will argue that both conceive knowledge as a reaction to an independent or pre-established ‘other’, which is either embraced or rejected. For these two standpoints, to be or not to be is still the question. But scepticism rises above this dichotomy and regards being itself as a misconception. To claim, categorically, that something *is* or *is not* is to construe identity and difference as fixed categories, falling outside the infinity of thought. For the sceptic, to be absolutely free one must abolish difference itself, and hence the very idea of predication—‘because the “different” is just this, not to be in possession of [oneself], but to have [one’s] essential being only in an other.’⁶⁶

What is achieved through scepticism is the recognition that reality ‘contains no permanent element, and must vanish before thought.’⁶⁷ But this vanishing, unlike the one endorsed by stoicism, is no longer merely replaced by the element of thought: the self does not merely *acknowledge* its own primacy, nor does it merely *detect* a general lack of correspondence between thought and reality. Scepticism is a mode of cognition ‘to which it does not *happen* that its truth and reality vanish without its knowing how, but which, in the certainty of its freedom, *makes* its other which it claims to be real, vanish.’⁶⁸ Scepticism amounts to the self-conscious denial of both the realist claim that truth is to be found in objectivity and the idealist claim that truth is to be found in thought.

Sceptical freedom is absolute because it is completely unbounded: since nothing is different from anything else, there is no good reason to claim this or that, to hold on to this truth instead of another. In this absolute dialectical unrest, every difference is dismissed as spurious and eliminated. But inasmuch as this elimination is itself grounded in a determinate judgement, it is once again dismissed, and unity is thrust back into difference. Caught in this endless cycle, scepticism is unable to provide consciousness with a stable cognitive perspective. Indeed, by refusing the possibility of a determinate version of reality, it is refusing the very possibility of cognition. Its cognitive horizon is reduced to a single, uninhabitable punctum, of which meaning,

⁶⁶PS, 124 / HW 3, 161

⁶⁷Ibid.

⁶⁸PS, 124 / HW 3, 160

identity and determination have all been expelled.⁶⁹ Therefore, scepticism can only maintain this negative attitude at the price of contradicting itself: it abolishes determination, but does so in a determinate way; it rejects being, but must be in order to do so; it dismisses everything stable and objective, but cannot help relapsing into stability and objectivity, if only to negate them over and over again.

This contradiction leads to the collapse of scepticism and to the endorsement of a new shape of consciousness. But this new stage is not yet the deliverance craved by both stoicism and scepticism. The ‘unhappy consciousness’, as Hegel calls it, amounts rather to the full acknowledgement of scepticism’s contradictions, and to a new and decisive turning point in the debate concerning the possibility of human freedom.

7.4 The Unhappy Consciousness

Scepticism entails a performative contradiction: the sceptic claims to have thrown off the shackles of meaning and determination, but this proclamation is itself meaningful and determinate. The paradox is unavoidable because determinacy lies at the very heart of human cognition and conditions all aspects of human life. The sheer fact of having a body, of reacting to sensuous stimuli, of taking up a specific space during a specific period of time—all of this forces the sceptic to return to objectivity and assent to the world he or she repudiates. Every qualification is a disqualification and every concrete reasoning leads back to identity, difference and predication.

This ceaseless interchange is scepticism’s first reaction to its own contradictions. Trapped between freedom and servitude, consciousness jumps from content to content, refusing to linger in any of them. But although it is aware of this inner restlessness, its conception of freedom is still one-sided. At this point, freedom still stands on one side, as something to be accomplished, while difference and determinacy stand on the other, as the obstacles preventing this accomplishment. But the experience of scepticism will put an end to this asymmetry. According to Hegel, the logical outcome of scepticism is the recognition that these two moments cannot be envisaged separately: thought’s absolute freedom is only conceivable as a reaction to the determinacy of the objective world; and the latter, in turn, only exists as a product of thought. Both moments are essential because they are both part of the same

⁶⁹Of course not all scepticism needs to be this radical. One might accept that knowledge is unattainable and still admit, for example, basic notions such as being, identity or difference. But Hegel is not interested in this kind of compromise. As always, his aim is to test the limits of this cognitive model and to bring out its inner contradictions. Moreover, his critique is conducted at a stage where there is no longer room for partial or incomplete forms of negation: the freedom pursued by self-consciousness is a radical kind of freedom, opposed to objectivity as such, and scepticism is therefore a radical kind of scepticism, opposed to objectivity as such.

thinking self. 'It is in fact *one* consciousness which contains within itself these two modes'⁷⁰, and for which the two are both necessary and incompatible. This acknowledgement leads to the abandonment of scepticism's dichotomic worldview and to the emergence of a new shape of consciousness, 'which *knows* that it is the dual consciousness of itself, as self-liberating, unchangeable, and self-identical, and as self-bewildering and self-perverting.'⁷¹

This new standpoint is appropriately named the 'unhappy consciousness', and the 'unhappiness' in question is twofold: it is due simultaneously to the continual urge to reconcile thought and reality and to the acknowledgement that such reconciliation is impossible. Consciousness is now fully aware that its remedy is also the cause of its evil, but cannot escape either: in attempting to free itself from the objective world, it only reiterates its captivity and the source of its unhappiness.

Like scepticism, the unhappy consciousness is initially torn between the pure universality of thought and the contingency of everyday life. One side is now referred to, in Parmenidean vein, as 'the simple Unchangeable' (*das einfache Unwandelbare*), whilst the other is defined, in Heraclitean vein, as 'the protean Changeable' (*das vielfache Wandelbare*).⁷² Yet these two aspects are not to be construed as independent or simply juxtaposed. The unhappy consciousness *is* and *is not* both sides of the division, and the aim of the ensuing dialectic is precisely to clarify this paradox. At the beginning of a long and complex exposition, Hegel sums up the dialectic's three main steps, corresponding to the different modes in which individuality can relate to universality:

1. Initially, self-consciousness sees itself as a mere individual opposed to an abstract and unattainable beyond. The Unchangeable is the essential moment of the relationship;
2. Afterwards, the Unchangeable is brought closer to the changeable moment by assuming an individual form. Both sides of the relationship become essential;
3. Finally, the Changeable finds itself within the Unchangeable and the two terms are reconciled.

At first, the meaning of these stages is rather obscure. As the exposition unfolds, however, Hegel's use of words such as 'elevation', 'devotion' or 'incarnation' evokes the specific context of religious experience. Through this change of vocabulary, the initial opposition between intellectual freedom and concrete individuality gives way to the opposition between a divine mode of existence and a secular one. But although this line of interpretation can prove very useful in decoding some of Hegel's most cryptic formulations, it must not overshadow the more general implications of his argument.

⁷⁰PS, 126 / HW 3, 163

⁷¹Ibid.

⁷²PS, 127 / HW 3, 164

In the previous sections, Hegel had already associated the phenomenological stages of stoicism and scepticism with their actual historical counterparts. In the present section, this association becomes even more explicit. As practically every commentator has recognized, the historical backdrop of the unhappy consciousness is the evolution of early Judaism, the rise of Christianity and its main developments throughout the Middle Ages. This historicized reading, encouraged by numerous textual hints, is further supported by the fact that this whole section amounts in many ways to a direct follow-up of Hegel's early theological investigations, 'transposed to a properly philosophical plane.'⁷³ In those early texts, the dilemma faced by the unhappy consciousness is already discussed in all its complexity, although with clearer references to its concrete historical context. Hegel's arguments revolve around the figures of Moses and Abraham, the life of Christ and the evolution of the Christian Church.⁷⁴

But how far should one rely on history to interpret Hegel's text? On the one hand, given its obvious historical echoes, it is tempting to suppose that the dialectic of the unhappy consciousness was originally conceived *ex historia*, and that the obscurity of Hegel's text is partly due to an effort to conceal the argument's historical skeleton, so to speak, and thereby preserve its universality. Moreover, it can hardly be denied that without previous knowledge of the historical and theological themes Hegel is alluding to, his exposition would be much harder to understand. On the other hand, it must be borne in mind that the *Phenomenology* is not a history of human consciousness, but a succession of cognitive models devised by consciousness to account for its own cognitive make-up. Accordingly, the unhappy consciousness is first and foremost a cognitive model, in the peculiar sense previously discussed,⁷⁵ which means that Hegel's description must be able to stand on its own, regardless of its historical echoes.

In the following pages, for the sake of expediency, I will focus simultaneously on the argument's logical and historical dimensions. Historically, the transition from scepticism to the unhappy consciousness is echoed by the decline of Greek culture and the rising influence of Jewish and Christian monotheism. For Hegel, the history of Greek philosophy is the history of subjectivity and its conquests. From Thales to Sextus Empiricus, the Greeks discovered and explored the infinite power of thought, redefining the scope and boundaries of self-consciousness. During this long *Ausbildung*, thought strived to 'tear itself away from what is sensuous' and to attain 'a totality beyond the sensuous and the imaginary'⁷⁶. But this one-sided movement was bound to be overcome: scepticism was both its climax and its dying hour, paving the way to a new philosophical age. The idea that thought could rise above

⁷³ Hyppolite, *Structure et Génèse de la Phénoménologie de l'Esprit*, 187

⁷⁴ See notably *Die Positivität der christlichen Religion* (1795–96) and *Der Geist des Christentums und sein Schicksal* (1798–1800).

⁷⁵ See Sect. 6.3.

⁷⁶ LHP, vol. 2, 451 / HW 19, 487

objectivity and overcome the contingency of the natural world was abandoned and the search for intellectual freedom gave way to a new form of servitude. The Greek world was replaced by the Jewish world and the focus on freedom and independence by a new fixation on the concrete aspects of human life.

In the *Phenomenology*, this transformation corresponds to the unhappy consciousness' first dialectical equation. Initially, self-consciousness regards the Unchangeable as the essential side of the relationship and places itself on the opposite side. Like the sceptical consciousness, the unhappy consciousness repudiates the contingency of its own existence, but unlike the sceptic it accepts this contingency as its own fate, or as the vainness to which life is naturally doomed. This first attitude brings to mind the opening lines of the *Ecclesiastes*, or the wretchedness endured by Job. It also replicates the servitude described in the master-slave dialectic, but on a different scale: slavery is now the self's natural condition, and the master is now divine and almighty.

Yet the unhappy consciousness is by definition both sides of the equation, which means that neither the Changeable nor the Unchangeable can be cast aside. Although the self is at first a merely contingent moment, it is immediately conscious that its essence lies elsewhere, namely in the Unchangeable. Therefore, it is led to reach out and claim possession of it. But since the latter is by definition unattainable, it can only reach the divine element through worship and adoration, that is, through the continual admission of its own servitude. Its life is therefore an eternal blend of hope and wretchedness, devotion and humiliation.

Despite the sinuosity of Hegel's exposition, the problem is really a simple one: the unhappy consciousness is bound to acknowledge the simultaneous validity of two self-excluding attitudes. If the immutable side is placed at a finite distance, its immutability disappears, and self-consciousness returns to the naïve realism of ordinary consciousness; if the mutable side passes over into the immutable one, self-consciousness returns to scepticism and to the contradiction that led to its present condition. The only way out of this paradox is to devise a form of relationship that preserves the independence of the two extremes. The solution must allow for the subsistence of a contingent *here* and a necessary *beyond*, but in such a way that both can be recognized as different sides of the same self-consciousness.

The search for this solution leads to the second equation, which Hegel associates with Christianity. The Christian concept of incarnation is offered as the answer to the Jewish dilemma and the figure of Christ is brought in to unite both sides of the antithesis. Through Christ, the 'embodied Unchangeable', the immutable side approaches the mutable one without relinquishing its immutability—for Christ is *simultaneously* Father and Son, divine and human, eternal and temporal. With this transition, the first Unchangeable, which was merely 'an alien Being who passes judgement on the particular individual', is replaced by a second Unchangeable, 'a form of individuality into which the entire mode of existence passes.'⁷⁷

⁷⁷PS, 128 / HW 3, 165

This second solution seems to accommodate the contradictions of the previous one: the concept of incarnation manages to overcome both the pure contradiction of scepticism and the simple self-coincidence of ordinary consciousness; in it, both terms come into contact with one another without sacrificing their independence. However, experience will show that this compromise is once again untenable. On closer inspection, Christianity's embodied divinity is just as inaccessible as Judaism's abstract one, for although the Christian is granted the possibility of experiencing the Unchangeable in its concrete form, the object of this experience is a sensuous object, and hence a changeable one. When the Unchangeable 'is sought as a particular individual, it is not a universal individuality in the form of thought, not a Notion, but an individual in the form of an object, or an actual individual'. Therefore, the body of Christ is really only 'an object of immediate sense-certainty, and for that very reason only something that has already vanished.'⁷⁸

Like all physical bodies, the body of Christ only displays its bodily nature. It confronts the observer with the silent obstinacy of sensuous objects, bound to a specific form and appearance. The idea that such a body is simultaneously a living image of eternity is something merely added to it—a mere label, irreducible to its physical presence. And although self-consciousness is attached to this label, it is forced to acknowledge its arbitrariness: it can neither explain why this particular body is an immutable essence nor why such an essence should be compelled to assume a bodily form. Both ideas are brought over from outside of the sensuous realm. They are the products of a new beyond, which is just as unattainable as the first one.

This new dialectic leads self-consciousness back to its previous dilemma: if incarnation is simply an idea, the Unchangeable remains distant and unattainable; but if divinity is successfully identified with individuality, its immutability is eliminated and Christ is reduced to a historical figure. To escape this contradiction, the unhappy consciousness moves on to a third and final solution, which is again evocative of the master-slave dialectic, and particularly of the slave's final emancipation. Since the previous attempts to set up the Unchangeable as the essential element have failed, the unhappy consciousness returns to the changeable element in search of deliverance. In a way, this new attitude is an acknowledgement of defeat: unable to ground its restlessness in a stable or convincing form of transcendence, consciousness returns to the stronghold of sensuous existence. From now on, its life is to consist in work and the enjoyment of earthly goods. The spiritual heights of divinity are replaced by the concrete depths of sensation: for here, at least, the fruits of one's labour are truly one's own and the objects of enjoyment are within reach.

Unlike the previous modes of self-consciousness, however, this return to the sensuous world is not immediate or one-sided. Objectivity is no longer merely the negative of self-consciousness, destined to be consumed or nullified, but 'an *actuality broken in two*, which is only from one aspect intrinsically null, but from another

⁷⁸PS, 131f. / HW 3, 169

aspect is also a sanctified *world*.⁷⁹ Conversely, the Unchangeable, which was formerly confined to a single individual, is now extended to the whole of objectivity. The unhappy consciousness sets about transforming the natural world and reaping the benefits of this transformation, but that is only possible because the Unchangeable has offered itself up as the basis for this activity. Nature itself is now the embodiment of the divine element and Christianity assumes a pantheistic dimension.

While the stoic and the sceptic proclaimed their independence vis-à-vis objectivity, the unhappy consciousness has come to see the world as a divine gift, for which it must grateful. But its humility, at this point, is not yet complete. Self-identity is still grounded in a personal core of self-certainty, which the self refuses to hand over to the Unchangeable. Consciousness still views itself 'as this particular individual, and does not let itself be deceived by its own seeming renunciation, for the truth of the matter is that it has *not* renounced itself.'⁸⁰ The fruits of its labour may well be provided by an alien source, but it can still use them up and enjoy them. Their divine provenance does not compromise their actual value, nor the pleasure they afford.

Though consciousness renounces the *show* [*Schein*] of satisfying its feeling of self [*Selbstgefühl*], it obtains the *actual* satisfaction of it; for it *has been* desire, work and enjoyment; As consciousness it has *willed, acted and enjoyed*. Similarly, its *giving of thanks*, in which it acknowledges the other extreme as the essential Being and counts itself nothing, is its *own* act which counterbalances the action of the other extreme, and meets the self-sacrificing beneficence with a *like* action.⁸¹

In other words, consciousness is willing to acknowledge the superiority of the universal element, but it refuses to give up its own individuality. This self-attachment revives the self-division maintained in the previous stages and constitutes a new source of unhappiness, which calls for a new dialectical compromise. In the remainder of the chapter, self-consciousness will be led to negate its previous negation, to alienate its own individuality and to acknowledge once more the absolute dominion of universality. But inasmuch as this negation is a *determinate* negation, this new alienation is different from the first one. Instead of reducing itself to a simple nothingness, immediately taken over by otherness, the unhappy consciousness will attempt to *find itself in its other* and turn its self-division into a truly reciprocal exchange.

We have seen this same movement play out in the beginning of the self-consciousness chapter, in the dialectic of recognition and again in the master-slave dialectic. In all of these cases, Hegel insists on the fundamental difference between an abstract or immediate form of self-consciousness (the simple tautology 'I am I') and a concrete or mediated form of self-consciousness (the 'identity of identity and difference', or the unity of selfhood and otherness). As he argues in different ways

⁷⁹PS, 133 / HW 3, 170f.

⁸⁰PS, 134 / HW 3, 171

⁸¹PS, 134 / HW 3, 172

throughout the *Phenomenology*, self-identity is not something given, but the result of a dialectical exchange. In order to become concrete, selfhood must give up its initial self-certainty, lose itself in otherness and return from otherness to itself—no longer as a simple ‘I’, but as a self-moving ‘I’, made up of identity and difference.

The same is now valid for the unhappy consciousness, although the ‘other’ in which it must lose itself is no longer another individual, or another self-consciousness, but a divine Being, or universality as such. Let us follow Hegel’s description of this final renunciation. Firstly, personal freedom is counted as vain and relinquished. Consciousness gives up its pretence of autonomy, forfeits the authorship of its actions and places itself at the mercy of an alien will. Secondly, life itself is divested of its former meaning and reduced to a purely physiological process. The ‘petty actions’ and ‘animal functions’ which used to be performed automatically are now the self’s primary focus, and a source of embarrassment and shame. Thirdly, personal property and the fruits of labour are likewise rejected. Enjoyment and pleasure are denounced as sinful and replaced by fasting and abstinence. In a word, self-consciousness embraces the ‘ascetic ideal’ that Nietzsche will later revile.

Furthermore, this radical renunciation also entails a profound cognitive transformation. Apart from forfeiting the right to choose and to carry out its choices, consciousness also gives up its quest for a meaningful existence and sets about ‘practising what it does not understand.’⁸² This aspect is extremely interesting, and would certainly profit from a more detailed exposition. What seems to be here at stake is not simply the observance of sacred rites or traditions, but also the adherence to religious dogmas, that is, the acceptance of truths whose validity no longer depends on one’s rational assent. And this new demand, even more than the previous ones, represents a complete break with the previous forms of self-consciousness. Whereas earlier the subjective moment had tried to obliterate, dominate or at least influence the objective moment, the notion of dogma entails an inversion of this pattern. Henceforth there is nothing left for the subject to choose or decide, no criterion with which to ground its perspective. Self-renunciation culminates in a pure leap of faith, which demands unconditional acceptance.

This point must be clearly spelled out: by ‘practising what it does not understand’, consciousness is not simply acknowledging its own ignorance, or surrendering to a higher epistemic authority. If that were so, the *incomprehensible* would amount to *what has not been comprehended*, but might come to be so. Consciousness’ self-alienation would remain incomplete, spoiled by the hope of acquiring a more thorough cognitive perspective. At this point, however, the dialectical resolution pursued by self-consciousness requires something like a *credo quia absurdum*: an acceptance of the incomprehensible *qua incomprehensible*.⁸³

⁸²PS, 137 / HW 3, 175

⁸³The Christian dogma of transubstantiation, for instance, can be accepted or denied, but it is not open to refutation. And it cannot be refuted precisely because it does not rest upon a logical argument or an empirical observation. Its truth hinges solely on the divine authority of its source.

Whereas stoicism, scepticism and the unhappy consciousness were all bent on 'explaining away' the self's inner contradictions, their efforts only gave rise to more contradictions. Their intrinsic faith in rationality has been identified as the primary source of self-division and dismissed as vain. Consciousness is now aware that its unhappiness cannot be overcome through reason or self-certainty. In order to rise to a new cognitive standpoint, it must give in to complete irrationality and embrace the mystery of faith.⁸⁴

Through this radical renunciation, the subjective core retained by the previous solutions is finally eliminated, and so is consciousness' self-alienation. Self-negation is finally revealed as a determinate negation and the self can finally place itself on both sides of the relationship. Although Hegel's account of this new *volte-face* is extremely concise, its general logic is by now familiar: just as self-consciousness, in its first appearance, has recognized itself as the author of the 'inner world',⁸⁵ and just as the slave, in the master-slave dialectic, has recognized itself as the 'author' of the natural world⁸⁶, so too does the unhappy consciousness now emerge as the source of its own transcendence. It is now aware, for the first time, that its essence is also its object, and that its object is also its essence. And this realization coincides with Hegel's conception of Reason, the next stage in the progression, defined as 'the certainty that consciousness, in its particular individuality, has being absolutely *in itself*, or is all reality.'⁸⁷

In the stage of Reason, the relationship between self and world is no longer disjunctive. Previously, consciousness had to settle for one of two opposite moments, both promising an independence they failed to deliver, but the advent of Reason marks the end of this dualism. In this new cognitive stage, consciousness and self-consciousness are finally brought together. Thought is no longer divorced from actuality and the world is both an idea and a lasting presence.

⁸⁴This attitude is echoed by the Christian condemnation of intellectual or philosophical hubris, voiced in different ways throughout the Holy Scriptures. Recall, for example, Christ's intent to 'destroy the wisdom of the wise' and 'bring to nothing the understanding of the prudent' (1 Cor 1: 19); or Isaiah's famous injunction: 'if you do not believe, you will not understand' (Isa 7:9), which proclaims the absolute subordination of reason to the authority of faith. To be sure, one might conceive the relationship between faith and reason in a more nuanced manner and envisage different ways of reconciling these two cognitive modes, namely a) by construing philosophical or scientific knowledge as a 'natural' or 'secular' translation of the immediate knowledge granted by faith (Augustine's *fides quaerens intellectum*); b) by regarding science and faith as different kinds of knowledge, directed at different kinds of objects (see Aquinas' *De Veritate* 14, 9); or c) by reconciling both terms through the pursuit of a philosophically or scientifically informed faith (see Leibniz, *Die philosophischen Schriften*, vol. 5, 476f.). In the present stage, however, philosophical insight and religious faith *must* appear as self-excluding, for they have both been driven to their most radical form. Since consciousness was led to embrace increasingly radical forms of subjectivism, its unhappiness can only be overcome by an equally radical form of self-alienation.

⁸⁵See Sect. 7.1.

⁸⁶See Sect. 8.1.

⁸⁷PS, 138 / HW 3, 177

Chapter 8

Towards Freedom



Abstract This chapter is centred on the *Phenomenology of Spirit*'s Reason and Spirit sections, where the three main antitheses that fuel Hegel's progression (subject-object, individual-collective, humanity-divinity) are brought into play, and retraces the evolution of the conception of freedom from the passive naturalism of 'observing reason' to the absolute self-derision of comedy. Section 8.1 discusses the dialectic of Reason and the recognition of Spirit (*Geist*) as a concrete, socially mediated and historically situated mode of cognition. Section 8.2 argues that the notion of Spirit turns freedom into an ethical and political ideal, transferring the *Phenomenology*'s educational programme to a social and historical context. Finally, Sect. 8.3 analyses Hegel's views on tragedy of comedy. These stages are particularly important for the present study, not only due to Hegel's explicit engagement with Socrates and Plato, but also because of their apparent inconclusiveness—the same inconclusiveness retained by Plato's critical standpoint, which Hegel aims to overcome.

Keywords Hegel · *Phenomenology of Spirit* · Freedom · Reason · *Geist* · Tragedy · Comedy

From the stage of self-consciousness onwards, Hegel's progression is fuelled by three fundamental antitheses. First, there is the general opposition between subjectivity and objectivity, or between reality *in itself* and reality *for consciousness*. This was the contrast that launched the progression and the reason why sense-certainty was abandoned. Unable to reconcile the spheres of selfhood and otherness, ordinary consciousness was led to transcend its initial standpoint and to travel from stage to stage, in search of its lost unity.

Second, there is the more specific, but no less decisive opposition between subjectivity and intersubjectivity, or between individual selfhood and a socially mediated identity. This theme was first made explicit in the self-consciousness chapter, as the subject of desire came to recognize its constitutive dependence on another desiring subject, and this dependence evolved into a mortal struggle, which

was then replaced by the dialectical exchange between a master and a slave. After this dispute, however, the theme of sociality was seemingly set aside, in favour of a new focus on individuality and transcendence.

This focus on transcendence brings us to the third main antithesis at play in the progression. With the experience of scepticism, the cognitive approach that had dominated the preceding stages was exhausted and a new world opened up before consciousness. Thus far, the progression had been confined to a horizontal plane, inhabited by human subjects and natural objects, but the unhappy consciousness transferred the debate to a vertical plane and introduced for the first time the idea of God. What thus came to light was a transcendent yearning, irreducible to the finite element of human life. The regional sphere of existence was articulated with the universal sphere of divinity, leading to a series of complex, but equally unstable dialectical equations, and this mounting tension was finally resolved with the emergence of Reason—but not altogether eliminated.

From now on, all of these elements will be revived, developed and revealed as the subplots of a wider overarching narrative. The stage of Reason, brought about through the immediate synthesis of consciousness and self-consciousness, is the meeting point of the first two antitheses. Its task is therefore a double task: on the one hand, it must be able to harmonize the conflicting demands of subjectivity and objectivity, and to reveal human cognition as a truly dialectical phenomenon; on the other hand, it must move beyond the individual or dualistic stances of the previous models and honour Hegel's definition of human consciousness as an inherently social phenomenon: the 'I' that is 'We' and 'We' that is 'I'. The articulation of these two movements will generate a series of increasingly sophisticated shapes of consciousness, which will eventually lead to the stage of Spirit, where Reason is finally raised to truth. This new phenomenological conquest will then reawaken the divine element left dormant in the previous stages, raise the progression to a new degree of complexity and set out its final itinerary.

8.1 Reason and Spirit

Throughout the previous stages, self-consciousness was at war with the world, forced either to negate objectivity in order to assert its independence or to give up its independence and surrender to objectivity. To break this cycle, the unhappy consciousness relinquished its usual self-confidence and placed itself at the mercy of a transcendent Being. As a result, its relation to reality turned round into a positive relation. Instead of retreating once more into selfhood, it came to find itself in the world it had repudiated and to feel at home in it. It came to realize that 'what *is*, or the in-itself, only *is* in so far as it is *for* consciousness, and what is *for* consciousness is also *in itself*.'¹

¹PS, 140f. / HW 3, 180

The acknowledgement of this reciprocity is the gateway to a new understanding of reality. But the full extent of this transformation is not immediately apparent. As Reason comes on the scene, it is still an immediate result. Consciousness has overcome its former self-division and is now ‘certain of being all reality’, but this certainty is at first nothing more than a simple form of idealism. For consciousness, reality is now *its own* reality, that is, a thoroughly *phenomenal* realm whose meaning is provided by Reason. Hegel is alluding, once again, to Kant and reiterating the criticism raised in the introduction and in the transition from the dialectic of consciousness to the dialectic of self-consciousness. If knowledge lies in the unity of apperception, the object of knowledge must be either thoroughly phenomenal, and hence unreal, or placed outside the phenomenal realm, and hence unknowable. This idealism is ‘empty’, Hegel argues, because it is stuck in this contradiction. While Reason provides the *form* of reality, its categories must be ‘filled’ with an objective *content*. But since the latter is brought in from the outside, from the noumenal world, it is by definition unknowable. ‘The pure reason of this idealism, in order to reach this *other* which is *essential* to it, . . . but which it does not have within it, is therefore thrown back by its own self on to that knowing which is *not* a knowing of what is true.’² It is forced to move back and forth between subjectivity and an independent objective beyond, be it Kant’s thing in itself or Fichte’s extraneous *Anstoß*.

To overcome this impasse, consciousness must recognize that the ‘I’ of idealism is still an abstraction. Instead of claiming *in abstracto* that ‘everything is its own’, it must set out to find itself in the objective world and to fill out its initial emptiness with the richness of natural existence. This movement is expounded at length in the dialectic of ‘observing Reason’, which amounts to a more sophisticated version of the dialectic of consciousness. In this new stage, to know the world is to know oneself, and to know oneself is to know the world. Contrary to what happened in the stages of sense-certainty and perception, objectivity is no longer an ‘other’ existing in its own right. Whereas ‘the perception and *experience* of various aspects of the Thing were something that merely *happened* to consciousness, . . . here, consciousness *makes its own* observations and experiments.’³ Its aim is to establish and bring forth the world’s intrinsic rationality, and to find itself therein as *actual* or *embodied* Reason. This project entails a long and painstaking journey, during which consciousness comes to recognize itself as an extraordinarily complex entity, embodied and articulated in a myriad of different ways. This journey is divided into three main stages, devoted to the observation of nature, the observation of self-consciousness and the observation of their mutual connection.

The first stage is in many ways the embryo of the *Encyclopaedia*’s Philosophy of Nature. This segment is fuelled by the continual tension between Reason’s theoretical constructs and the contingency of natural phenomena. While the phenomenological subject attempts to subsume reality under universal patterns and laws, the natural world keeps eluding these attempts, proving more complex and

²PS, 145 / HW 3, 184f.

³PS, 145 / HW 3, 185

unpredictable than initially supposed. The limitations of the observational approach are felt at first in connection with the study of physics and geology, and aggravated with the study of biology and zoology. In its quest for the universal principles underlying the morphology and behaviour of living beings, Reason sets out to test a series of explanatory models, leading to complex debates about the influence of environmental conditions on living organisms, the nature and scope of internal and external teleology, the joint action of sensibility, irritability and reproduction. In the end, however, none of these approaches is capable of bridging the gap between the universality of thought and the particularity of the natural world. Consciousness is left with a series of 'interesting connections', whose conflicting insights cannot be raised to a coherent rational system.

This failure leads Reason to turn inward and focus on the laws that govern the human mind. The second and third stages of 'observing Reason' deal with the observation of self-consciousness as regards both a) the nature and purpose of its thoughts and actions (observational and social psychology) and b) the meaning of its bodily expression (physiognomy and phrenology). In the first case, the debate is centred on the issue of how and to what extent human thoughts and actions are influenced by one's natural and social environment. Hegel argues that 'influence' (*Einfluß*) is a vague and inconsistent notion, and that human behaviour is ultimately irreducible to a constellation of natural and social circumstances. In the second case, consciousness turns once more to objectivity, in an attempt to harmonize its psychological insights with the natural laws uncovered in the first stage. Hegel explicitly addresses two pseudosciences that became popular during the eighteenth century and proposed to draw psychological conclusions from the observation of human beings' physical features—in the case of physiognomy, by studying people's facial features; in the case of phrenology, the shape of their skulls. Again, Hegel's refutation consists in exposing the absence of a necessary link between physical features and mental dispositions. Although Reason is indeed as much 'mental' as it is 'real', it cannot be placed inside a muscle, a bone or a brain fibre.

The failure to unify thought and being leads consciousness to realize that humans are not 'things' to be 'observed'. Human thoughts and actions are dependent on a principle of self-determination that eludes observation, and this fundamental insight alters the entire focus of the enquiry: henceforth the rational subject will no longer be defined as an *observer*, but as an *agent*. This new stage can be interpreted as a more sophisticated version of the dialectic of self-consciousness, where the self 'no longer aims to *find* itself *immediately*, but to produce itself by its own activity.'⁴ What is now at stake, therefore, is no longer the subject's intrinsic nature, but its self-actualization. And since the latter is to take place in a world inhabited by other self-actualizing subjects, the theme of social recognition is explicitly reintroduced.

Rational self-consciousness starts out as an egoistic standpoint, fuelled by an individual urge for satisfaction. This characterization evokes the dialectic of desire, but with an important difference: unlike the desiring consciousness, intent on annihilating the objective world and reaffirming its absolute sovereignty, this form

⁴PS, 209 / HW 3, 261

of consciousness is aware that the objective world is its own world and that its essence is to be found therein. Yet despite this reciprocity, the self's outlook on reality is just as selfish as before: the world is still conceived in a purely instrumental manner, as a vehicle for the attainment of enjoyment and satisfaction. Hegel's description is also reminiscent of the master-slave dialectic in that it focuses not on the pleasure of eating or drinking, but on social and sexual domination. The pleasure sought at this stage also depends on the interaction with another self-consciousness. However, this interaction is not yet the mediated unity of social recognition, but the immediate unity of lust and gratification. The other still amounts to a 'negative essence', 'devoid of reality and content.'⁵

Since the object of pleasure is devoid of reality, each act of gratification is the source of a new urge for gratification. This inescapable cycle is not so much gratifying as self-defeating. Its true result is not a free, fulfilling life, but a life of obedience, whose permanent restlessness is eventually cut short by death. It is the life of *πλεονεξία* advocated by Callicles and exposed by Socrates as a form of slavery. Faced with this gloomy prospect, self-consciousness gives up its initial egoism, widens the scope of its outlook and develops a philanthropic interest in the welfare of mankind. Hedonism is replaced by the 'law of the heart', which asserts human beings' universal right to happiness and self-fulfilment. While personal pleasure was essentially base, the law of the heart is noble and high-minded. It represents the first phenomenological attempt to formulate a universal moral imperative.

For this new rational agent, dissatisfaction and unhappiness are not unavoidable, but the product of an unjust world, ruled by unjust laws and institutions. Reason's new mission is therefore to *reform* the world and bring out humanity's original goodness. But this project is once again discredited, for two main reasons. First, inasmuch as the law of the heart is still the law of *one* heart, abstractly proclaimed as the law of *all* hearts, its reforming task is doomed to failure. The philanthropist is either bound to recognize the naïvety of his efforts and to give up his mission or to lapse into fanaticism and impose his righteousness on the whole world. Second, since the law of the heart is a pure moral imperative, its actualization amounts to a self-betrayal. Hegel makes this point in the following manner: 'the law of the heart, through its very realization, ceases to be a law of the *heart*. For in its realization it receives the form of an [affirmative] *being*, and is now a universal *power* . . . , so that the individual, by setting up his own ordinance [*Ordnung*], no longer finds it to be his own.'⁶ In other words, once the reformer's abstract righteousness is actualized, the heart is deprived of its moral authority and becomes entangled in the worldly dynamic it was supposed to transcend. As a result, self-consciousness' reforming efforts lead to social estrangement—inasmuch as other individuals reject the reformer's dogmatism, and choose to obey the laws of their own hearts—and self-estrangement—inasmuch as the reformer himself, torn between moral 'activism' and moral 'puritanism', fails to bring together the model's conflicting claims.

⁵PS, 219 / HW 3, 272

⁶PS, 223 / HW 3, 277

The previous models collapsed as a result of Reason's inability to turn its individual aspirations into universally valid goals: while the individualism of pleasure was crushed by the universal law of necessity, the law of the heart failed to harmonize its universal ideals with the actual diversity of other people's interests. But although these models have proven untenable, they represent an initial departure from Reason's initial egoism. In the last shape of 'active Reason', this movement away from the individual and towards the universal becomes consciousness' central concern. In contrast both to the individualism of pleasure and to the individual universalism of philanthropy, what emerges is a truly *universal universalism*, in which individuals *qua individuals* no longer play any role.

This new standpoint—the standpoint of virtue—is really the standpoint of pleasure *in reverse*: while the latter was selfish and self-centred, the former is entirely altruistic; while the latter sought to achieve satisfaction at the expense of universality, the former must be accomplished at the expense of individuality. Virtue consists in a selfless pursuit of the common good, culminating in a universal concord reminiscent of Rousseau's *volonté générale*.⁷ The very idea of particularization is rejected and every personal aim is seen as illegitimate and ultimately perverse.

In short, virtue's emphasis on universality is just as abstract as hedonism's emphasis on individuality. The contradiction in question is also similar to the one faced previously by the stoic self-consciousness: since virtue's only mandate is to abolish particularity, its ideal is an empty one, i.e. universality *as such*. But human beings cannot live or act on such a mandate. Their thoughts and actions require a specific content, and are thus bound to relapse into particularity. As Charles Taylor puts it 'man cannot suppress his particularity, and act just as a vehicle of the universal; for he cannot simply act on the motive of conforming his action to universal maxims, setting aside all other motives. To refuse all other forms of action is to do nothing.'⁸

Virtue's immobilism is opposed to what Hegel calls the 'way of the world' (*der Weltlauf*), that is, the acknowledgement that human motivations are in fact intrinsically selfish, and that the common good is really only the global sum of countless individual interests. This atomistic view, characteristic of modern political liberalism, is brought in to contradict virtue's sterile utopianism. But Hegel's criticism is just as much directed at the virtuous self as it is at the 'way of the world'. Its aim is to overcome the antithesis expressed by their opposition and to reveal reality as the undivided unity of particularity and universality. The 'fatuous rhetoric' of virtue is bound to be 'dropped like a discarded cloak',⁹ but the individuality of the 'way of the

⁷ See notably Rousseau, *Du Contrat Social*, 277.

⁸ Taylor, *Hegel*, 166. As most commentators have recognized, the dialectic of Active Reason evokes the period of Enlightenment, and particularly Rousseau's moral and political philosophy. While the 'law of the heart' mirrors his belief in the fundamental goodness of human nature and his denouncement of the corrupting effects of culture and civilization, the dialectic of 'virtue' echoes his views on democracy and popular sovereignty.

⁹ PS, 234 / HW 3, 290

world' must also be abandoned: when it 'acts in its own interest, it simply does not know what it is doing; and when it avers that everyone acts in his own interest, it is merely asserting that no one knows what action is.'¹⁰

The fate of 'active Reason' is to move beyond the empty idealism of virtue and the resigned selfishness of liberalism. But this can only come about through the sublation of the long-standing antithesis between self-determination and self-actualization. Accordingly, in Reason's third and last main section, the subject's task is no longer to determine whether his or her actions agree with an underlying essence or criterion—be it the selflessness of 'virtue' or the selfishness of the 'way of the world'. Individuality has become 'intrinsically real', which means that the end of one's actions must now be looked for in the actions themselves. There is no longer an independent self presiding over the world, nor an independent world presiding over the self. The role of self-consciousness is henceforth to display through its own actions the unity of reason and reality. Self-identity is equated with *self-expression* and subjects become creators or producers, whose true identity lies in their work.

The ensuing dialectic is once again reminiscent of the master-slave scenario. Just as the slave had 'put himself in his work', and 'shaped the Thing in his own image', so too will the rational self express its identity through its activity. Yet just like before, the coincidence of selfhood and self-expression is not as straightforward as it appears. For one thing, since the product of one's work is usually finite and contingent (a text, a painting, a house, etc.), it is by definition incapable of matching the infinite complexity of self-consciousness. Furthermore, since this product has a fixed shape, it can only capture the nature of self-consciousness at the time of its creation—or not even that, for the creative impulse may be shorter than the actual creation, and the work may already be outdated in the moment of completion. These and other contradictions open up a new gap between self-identity and self-expression. They suggest that what one *is* is irreducible to what one *does*.

Moreover, Hegel notes that the nature of one's work is not determined solely by one's creative efforts. In a social environment, its meaning is shaped by other people's judgement, and the latter influences the creator in return. Every act of self-expression entails an intersubjective exchange that alters the very identity one is supposed to be expressing. This co-implication is described in terms of the dialectical opposition between one's commitment to 'the heart of the matter' (*die Sache selbst*)—a selfless immersion into the specific challenges and requirements of one's work—and an undeclared concern for success and recognition—which turns one's professional commitment into 'deceit' (*Betrug*).¹¹ What one expresses through one's work is not simply what one is, but also how one wants to be perceived.

¹⁰PS, 235 / HW 3, 291

¹¹See PS, 237ff / HW 3, 294ff. Robert Solomon gives the example of 'a scholar laboring for years on a study of a philosopher who has been dead for 150 years; on the one hand, no task could be more disinterested and free from the suggestion of personal gain; on the other hand, no enterprise would be more personally cautious since, for the scholar—as Hegel indelicately puts it—"the most important thing is his own vanity."' (Solomon, *In the Spirit of Hegel*, 517)

After establishing that the product of one's work is not an adequate vehicle for self-expression, Hegel goes on to argue that this inadequacy is ultimately due to the limited scope of the dialectic. Individuals are incapable of overcoming the contradiction between earnestness and deceit precisely because they still think in individual terms. Granted, the dialectic has already shown that self-expression is not a private process, but one that is influenced or affected by public opinion. But this is not enough. Reason must move beyond this superficial insight and reveal the whole dynamic of self-expression as an intrinsically social process.

Each individual, in seeking to express himself or herself through work, is confronted by other individuals who seek to do the same. What we have, then, is not one rational agent opposed to others *in general*, but a highly complex web of self-expressing agents, all of whom aim simultaneously at the 'heart of the matter' and at social recognition. But Hegel's point is not simply that one's work is influenced (or perverted) by social exposure. He also claims that the work of each individual is a *condition*, in a more or less direct manner, of other people's work. And this is made clear by observing how one's interest in a specific subject is often the starting point of other people's interest on the same subject, or how one's opinions often derive from the opinions of others. These observations alter the nature of the 'heart of the matter' and reveal each individual's activity as the contribution to a *universally shared activity*. If, on the one hand, what is essential in work is the expression of each agent's particularity, on the other hand 'it is learned by all parties that they *all* are affected, and regard themselves as invited [to participate].'¹²

Individual activity is thus but a moment within a wider social web. Yet this conclusion does not exclude the possibility of a return to the 'way of the world'—that is, the dependence of each on all may simply result in a global mosaic of self-interested standpoints, led out of necessity to collaborate with each other. This is not, however, the kind of interdependence Hegel has in mind. He aims at a stronger universal bond, irreducible to the atomistic logic of the previous standpoints. The 'heart of the matter' is such 'only as the action of *each* and *everyone*: the essence which is the essence of all beings, viz. *spiritual essence*.' It is the '*subject* in which there is individuality just as much as *qua* individual, or *qua this particular* individual, as *qua all* individuals.'¹³

This crucial sublation paves the way for the emergence of Spirit, and for the *Phenomenology's* second half. But before closing the net on individuality, Hegel dwells a little longer on the inner structure of this newly found standpoint. Since individual activity was revealed as the 'action of each and everyone', human actions must now be guided by universally valid laws, which can be acknowledged and shared by everyone. The remainder of the Reason chapter is dedicated to finding these laws, by turning to the realm of morality. In particular, Hegel focuses on the dichotomy between an individual conception of morality—*Moralität*, based on the

¹²PS, 251 / HW 3, 310. Translation modified.

¹³PS, 252 / HW 3, 310f.

Socratic-Kantian imperative of rational self-determination—and a social conception of morality—*Sittlichkeit*, or ‘ethical life’, based on a wider system of communally shared values. This key Hegelian theme, already present in earlier writings, is further developed in the *Encyclopaedia* and the *Philosophy of Right*, and its importance cannot be stressed enough. The idea that self-determination is from the very beginning a communal affair is the motor behind the transition from a philosophy of Reason to a philosophy of Spirit, and the key to understanding the articulation between the *Phenomenology*’s first and second halves.

The argument starts with Reason’s acknowledgement of the need for a morally grounded theory of action: if human actions are to rise above mere arbitrariness, human agents must be in possession of a universal criterion with which to distinguish what is right from what is wrong, what should be done from what should be avoided. Hegel’s primary interlocutor here is Kant, and particularly his *Metaphysics of Morals*. The issue can thus be translated in Kantian terms as the need to guide oneself by a ‘universal law of action’, which ‘must serve human will as its only principle.’¹⁴

The question then becomes how to determine this law, or rather how to establish a universal criterion with which to distinguish right from wrong. A possible answer is Kant’s so-called ‘universalization’ test—viz. his contention that the validity of a moral precept must be assessed not on empirical grounds, but by determining whether it can be successfully transformed into a universal maxim.¹⁵ For Hegel, this criterion is inconclusive and incapable of grounding a truly operative moral system. To illustrate his point, he turns to a brief discussion of the notion of property, aiming to show that the law that there should be private property and the law that there should *not* be private property can both be said to fulfil and to violate Kant’s formal criterion: when considered in isolation, they are both logically consistent (and hence ‘universalizable’), but in the concrete, when it comes to specific rules and regulations, they are easily shown to be contradictory (and hence ‘non-universalizable’). Hegel argues that this ambiguity extends to every other moral law and proceeds to announce (perhaps too swiftly)¹⁶ the *Aufhebung* of Reason’s last phenomenological resort.

More generally, Hegel argues that the standpoint of Reason, in its attempt to test the content of moral laws, is still confined to the sphere of *Moralität*. The principles that guide human actions are still construed as a series of commandments prescribing

¹⁴ *Kant’s gesammelte Schriften*, vol. 4, 402

¹⁵ See *ibid.*, 402, 421, 437ff.

¹⁶ Some commentators have denounced the superficiality of Hegel’s engagement, at this stage, with Kant’s moral philosophy. Indeed, it is not at all clear that Hegel’s passing discussion of the right to own property should result in a complete or convincing refutation of Kant’s universalization principle. For readings that are sympathetic to Hegel’s objection, in the *Phenomenology* and elsewhere, see Smith, *Hegel’s Critique of Liberalism*, 73ff.; Taylor, *Hegel*, 168ff., 371; or Walsh, *Hegelian Ethics*, chap. 4. For a discussion of Hegel’s engagement with Kantian morals, see Pippin, *Idealism as Modernism*, 104–128; Wood, *Hegel’s Ethical Thought*, 154–173; Stern, *Kantian Ethics*, chap. 8.

what ‘ought’ to be, which can be accepted or rejected according to individual rational judgements. The moral examiner places himself above human life, but his or her mandates are still ‘only a *willing* and *knowing* by this particular individual’,¹⁷ unilaterally raised to universality.

Unlike morality, *Sittlichkeit* is not abstract or arbitrary, but concrete and actual. It is not the result of a decree, but the native element of human life. As Hegel puts it, ethical laws simply *are*—they are *already there, already at work*, woven into one’s very identity, embodied from times immemorial by the institutions and traditions that shape human existence. Their content is not something to be posited or tested, but something to be acknowledged. Accordingly, if these laws ‘are supposed to be validated by *my* insight, then I have already denied their unshakeable, intrinsic being and regard them as something which, for me, is perhaps true, but also is perhaps not true.’¹⁸ This, for Hegel, is contradictory, since the ‘ethical disposition consists just in sticking steadfastly to what is right, and abstaining from all attempts to move or shake it, or derive it.’¹⁹ Now these and similar statements might be taken to mean that ‘ethical life’ is really about an unquestioning submission to the status quo. If that were the case, this new stage would surely amount to a step back rather than a step forward in consciousness’ phenomenological journey. But Hegel’s point, at this stage, is simply that *Sittlichkeit* is the necessary starting point for any serious consideration of human life. In order to be examined and judged, ethical values must first be acknowledged, which can only truly happen once Reason ceases to be regarded as an individual phenomenon, detached from its social and historical environment. ‘This is not to say that *Sittlichkeit* has not rules, or that the rules cannot be criticized or even rejected on rational reflection; but it is to insist that first there must be such rules and the practices in which they are embodied. Society is not, as in Hobbes and Rousseau, first of all a conference for the purpose of establishing the rules; society is that set of rules and practices. Their justification and legitimacy, the search for reasons, comes later, if at all.’²⁰

With the dawn of the ‘ethical world’, the atomism that presided over the previous stages is finally sublated and the subject sees itself in a new light. Consciousness, Self-Consciousness and Reason, falsely posited as concrete cognitive principles, are now unmasked as the abstract, pre-social and pre-historical semblances of what Hegel names Spirit, i.e. a *concrete, socially mediated and historically situated* self:

Spirit is thus self-supporting, absolute, real being. All previous shapes of consciousness are abstract forms of it. They result from Spirit analysing itself, distinguishing its moments, and dwelling for a while with each. This isolating of those moments presupposes Spirit itself and subsists therein.²¹

¹⁷ PS, 260 / HW 3, 320

¹⁸ PS, 261f. / HW 3, 322

¹⁹ PS, 262 / HW 3, 322

²⁰ Solomon, *In the Spirit of Hegel*, 534f.

²¹ PS, 264 / HW 3, 325

Every human self is a ‘spiritual’ self—this decisive acknowledgement is the result of a double movement, whose advances and setbacks were documented in the previous chapters. Consciousness has struggled to bring together, on the one hand, the spheres of subjectivity and objectivity, and on the other hand the spheres of subjectivity and intersubjectivity. Accordingly, Spirit is not a simple, but a double synthesis, or the simultaneous sublation of these two dialectical contrasts.

As regards the first one, recall that the notion of Reason was initially defined as consciousness’ ‘certainty of being all reality’. The antithesis between subject and object, maintained throughout the dialectics of consciousness and self-consciousness, was replaced by the dialectical unification of selfhood and otherness, but this new coincidence was not immediately understood. Confined at first to an empty form of idealism, Reason was led to move out into the world and to search for truth in the concrete element of natural existence. The transition to the stage Spirit is the result of this long and eventful journey. In Hegel’s words, ‘Reason is Spirit when its certainty of being all reality has been raised to truth, and it is conscious of itself as its own world, and of the world as itself.’²² While initially Spirit was consciousness that only *has* Reason, it has now been grasped as ‘Reason that *exists*, or as Reason that is *actual* in Spirit and is its world.’²³

As regards the second contrast, consciousness has hitherto attempted to combine its own identity and interests with the identity and interests of others, by means of different conceptions of personal and common good. But these efforts took the form of arbitrary principles, disconnected from the actual laws and traditions that shape the existing world. The acknowledgement of this disagreement led to the replacement of the moral framework adopted by Reason with the ethical framework introduced by Spirit, recognized at last as the underlying principle of human life. Whereas the dialectic of Reason was still about individual self-determination, Spirit is ‘the *ethical* life of a people [*Volk*], in so far as it is the *immediate truth*—the individual that is a world.’²⁴

Henceforth, the progression will no longer deal with individual shapes of consciousness, but with collective ones. This does not mean that the *Phenomenology* will now focus on crowds instead of individuals, or that it will offer a general exposition of world history. Hegel’s point is rather that the essence of human thoughts and actions cannot be grasped through an examination of their personal motivations. It calls for a wider analytical approach, able to capture both the specific nature of each individual self-consciousness and the universal element in which it is grounded. Since each individual is now a whole world, Spirit is likewise the dialectical unity of individuality and universality: it is ‘the unmoved solid *ground* and *starting-point* for the action of all’, but it is also ‘a fragmented being, . . . in which each accomplishes his own work, rends asunder the universal being, and takes from it his own share.’²⁵ The immanent cohabitation of these two moments will fuel the experience of Spirit, outlined in the following sections.

²²PS, 263 / HW 3, 324

²³PS, 265 / HW 3, 326

²⁴Ibid.

²⁵PS, 264 / HW 3, 325

8.2 The Trials of Spirit

Spirit is at first ‘true spirit’ and its original seat is the ethical world. In this new phenomenological setting, individuality and universality are immediately joined together. The private and the common good are not at variance with one another because each finds in the other its natural expression. The dialectic of *Sittlichkeit* draws on a theme that is dear to Hegel and to a whole generation of German thinkers. The harmony of the ethical world is illustrated by the harmony of the ancient Greek world.²⁶ The latter represents the first incarnation of the ‘ethical substance’: it is an ‘immaculate world’, ‘unsullied by any internal dissension’,²⁷ where individual desires are naturally aligned with communal goals.

But the ethical world is nonetheless a self-differentiated world. Spirit’s universal and individual moments are present in it in the form of two general laws: a divine law, associated with the private sphere of the family, and a human law, associated with the public sphere of politics. The first of these laws proceeds directly from the gods and regulates the most basic aspects of human existence. Its content is passed on from generation to generation, honoured by ancestral rites and ceremonies. The human law is upheld by the state and regulates the life of the polis. Here, individuals are not regarded as family members, but as citizens called upon to fulfil their duties towards the community.

Initially, these two laws complement each other. While the family’s main role is to educate and prepare the individual for the duties of citizenship, the state’s main role is to ensure that the citizen’s actions serve the interests of the community. At the end of the citizen’s life, however, this movement is reversed: his body is given back to the family, in order to be buried, and the divine law claims back its initial sovereignty. This cycle is harmonious because both laws have jurisdiction over different aspects of human life, thereby ensuring the maintenance of the ethical order. But what happens if they are called upon to judge the same human deed? This is the scenario envisaged in the next stage of the dialectic, as Hegel starts unpacking the ethical world’s inner contradictions.

The main reference for this new dialectical stage is Sophocles’ *Antigone*. The play’s plot is well-known: following the death of Polynices, pretender to the throne of Thebes, the city’s new ruler, Creon, condemns his actions as treasonous and decrees that his body remain unburied²⁸; Antigone, Polynices’ sister, defies this

²⁶ Hegel is here influenced by Herder’s, Schiller’s and Hölderlin’s views on the opposition between the lost innocence of Greek culture and the dividedness and disenchantment of the modern European world.

²⁷ PS, 278 / HW 3, 341

²⁸ Polynices and Eteocles, sons of Oedipus, had both been offered the throne of Thebes, but Eteocles managed to expel his brother from the city with the support of the Theban people. Polynices travelled to Argos, where he rounded up a small army, and launched a campaign to reconquer Thebes. In the ensuing battle, the two brothers ended up killing each other and Creon was crowned king of Thebes. In his eyes, Polynices is a traitor because he turned against his native city, and must not be given any burial rites.

order, buries her brother's body and is subsequently incarcerated. The conflict depicted in the play stems from the fact that both Antigone and Creon are legitimized by a different law and unwilling to change their position. Antigone represents the divine law and views her brother as a relative to be honoured; Creon represents the human law and views Polynices as a citizen to be punished. The conflict is insoluble because both parties are immersed in their natural roles and incapable of considering the other's point of view. Entrusted with an immediate ethical mandate, they have no choice but to affirm what they know to be right.

The ethical world has reached a deadlock. Since the law is still construed in an immediate or natural way, social roles are still merely given. They are one-sided stances thrust against each other and bound for destruction. The acknowledgement of this limitation reveals 'the germ of deterioration inherent in the beautiful harmony . . . of the ethical Spirit'²⁹ and precipitates the collapse of the Greek world. As the bond that unites the citizens is severed, the ethical community breaks down into a simple aggregate. Citizens are replaced by free-standing 'persons', committed to the pursuit of their own interests. This new scenario, illustrated by the juridical freedom of the Roman world, echoes the previous transition from the stage of 'virtue' to the 'way of the world'. Society is again reduced to the total sum of its members and the common good to a mosaic of particular interests.

While classical citizenship rested on a concrete conception of self-identity, personhood is a formal category. A person is no longer defined by its convictions or place within the community, but by the sheer fact of being a person. And since this is still an empty qualification, it must be filled up with external attributes. Social recognition becomes dependent on the acquisition of property, which leads in turn to a redefinition of personal freedom. Previously, the welfare of others was a condition of personal welfare; now, the latter is the primary element of the relationship. Freedom is reduced to a negative concept, that is, to the freedom of pursuing one's interests without the interference of others. And since this new prerogative is no longer grounded in a natural equilibrium, it must be upheld in an artificial manner, through the legal regulation of civil rights.

Each person is thus a self-interested, legally protected, property-owning individual. This definition is not Greek, but thoroughly modern. Historically, it leads as far as the advent of civil society and the rise of political and economic liberalism. But Hegel's primary aim is not to paint this wider historical picture. As always, he is interested in the model's immediate contradictions, and finds them at hand in the Roman world. His argument consists in showing that a society exclusively predicated on self-interest is bound to lead to inequality and despotism. The contradiction lies in the fact that the right to accumulate property, and hence to guarantee one's freedom, is ultimately self-defeating, as it leads to a disproportionate concentration of wealth. In a society dominated by the rich, political power is disconnected from the common good and the legal order gives way to an arbitrary one, where law and morality are outweighed by greed and violence.

²⁹PS, 289 / HW 3, 354. Translation modified.

Led from the harmony of the Greek world to the disharmony of the Roman world, Spirit must attempt to restore its self-integrity. The dialectic of personhood thus gives way to the dialectic of 'self-alienated Spirit', focused on the contradictions of modern European culture. Hegel's argument here consists once more in a series of variations on the relationship between the individual and the state. The first attempt at compromise is embodied by a 'noble consciousness' for whom the state is a fundamentally just institution, whose main concern is the prosperity of its subjects. Committed to this vision, the noble individual embraces a life of heroism that honours and strengthens the state. However, this abnegation revives the contradiction revealed in the previous model. As the state grows stronger, the distance separating it from the individual also increases and nobility gives way to increasingly radical forms of subjection. Eventually, the individual is forced to recognize the authority of an absolute ruler and to adopt a subservient and flattering attitude.

The transition from nobility to flattery, accompanied by suggestive allusions to French absolute monarchy, prompts the emergence of an 'ignoble consciousness' and a feeling of revolt. The state is now perceived as a despotic and arbitrary power, opposed to private welfare, and its wealth is no longer a public asset, but an instrument of domination. The noble spirit recognizes the vanity of its sacrifices but is materially dependent on the state and unable to break free. Instead, its integrity degenerates into a scornful cynicism, aimed at everything supposed to count as serious, important or noble. This new attitude is the opposite of the idealism expressed at the beginning of the dialectic. Spirit is no longer noble or ignoble, but the very consciousness of this opposition. Whenever something is deemed important or admirable, it points out its intrinsic baseness; whenever something is deemed base, it immediately praises its nobility. By reducing everything to its contrary, it celebrates the 'absolute perversion' of modern social life. Its only truth is a 'destructive judgement' aimed at the whole range of accepted norms and values, 'a nihilistic game' which 'overpowers everything.'³⁰

This 'disrupted consciousness' (*zerissenes Bewußtsein*) is not just faced with a contradictory world. 'As self-consciousness in revolt', 'it is aware of its disrupted state, and in thus knowing it has immediately risen above it.' Through the conscious dissolution of everything serious and stable, 'all content is turned into something negative', and what remains 'is merely the *pure "I" itself*, . . . this pure self-identity of self-consciousness that has returned to itself.'³¹ The next dialectical stage documents, therefore, a new return to self-consciousness, which can take two different forms. In the first one, Spirit proposes to find a refuge from the vanity of the outside world within itself, in the pure essentiality of thought. In the second, it refers back to a supersensible world, placed beyond subjectivity and objectivity. In the first case,

³⁰PS, 317 / HW 3, 386. Hegel's primary source here is Diderot's *Le neveu de Rameau*, which relays the dialogue between an 'honest', philosophically-minded narrator and the extravagant nephew of Jean-Baptiste Lully, whose incendiary views on morality, art and society lay bare the contradictions of modern bourgeois culture.

³¹PS, 321 / HW 3, 390

Spirit becomes ‘pure insight’ and endorses a thoroughly rationalist worldview; in the second case, Spirit becomes ‘faith’ and endorses a religious worldview. What follows is the dialectical confrontation of these two attitudes, epitomized by the historical conflict between Christianity and the revolutionary ideals of the Enlightenment, a debate that dominated Europe from the late seventeenth century up until Hegel’s lifetime.

As always, the dispute starts out as a simple antithesis, with each side rejecting the other. For the Christian, the Enlightenment’s unconditional faith in rationality is naïve: human reason, just like the natural world, is ultimately a divine gift, whose contradictions can only be solved by appealing to a higher cognitive plane. For the enlightened thinker, religion amounts to a series of mistakes and superstitions, imposed on the ignorant by the ‘the deceiving priesthood’ and ‘the oppressive despot’.³² The Enlightenment’s aim is to undo this domination, reawaken human beings’ rational instincts and bring them over to its own side of the barricade.

Hegel maintains, throughout this section, that the opposition between faith and reason is an abstract opposition, and that both sides have much more in common than they are willing to admit. But this is not yet the place for a reconciliation. The dialectic will focus, rather, on what Hegel views as the overall superficiality of the Enlightenment’s understanding both of religion and of its own rational project. Among the various accusations levelled against Christianity, Hegel singles out the denouncement of faith’s supposed unrealism (the idea that God is either a mental fabrication and/or a political deception), its contradictory relapse into empiricism (the worship of physical objects, the reliance on the Bible’s historical text) and its ‘foolish’ asceticism (the irrational renouncement to physical and natural enjoyment). All of these ideas reveal, rather, the Enlightenment’s constitutive inability to grasp the true essence of religious faith.

Afterwards, the exercise is reversed and the Enlightenment is called upon to offer its own perspective. Averse to all forms of transcendence, enlightened thinkers turn their focus to the natural world and confine reason to a fully secularized standpoint, whose main ideals are concrete objectivity and utility. *Das Nützliche*, the ‘useful’, becomes Spirit’s true essence, and with it the Enlightenment claims to have put an end to faith’s vacuous musings and ‘transplanted heaven to earth down below’.³³ This presumption will be tested in the third stage of the dialectic of self-alienated Spirit, where this ‘earthly heaven’ is transferred to the social realm and transformed into a concrete political doctrine.

The Enlightenment’s political worldview is brought in to correct the contradictions of the previous social models, where the individual and the state remained at odds with one another. In the Greek ethical order, private and communal interests were equated, but there was yet no room for self-determination. In the Roman legal order, the emphasis was placed on individual freedom, but the common good was

³²PS, 330 / HW 3, 401

³³PS, 355 / HW 3, 431. Translation modified.

thereby sacrificed. Now, the positive aspects of the two previous models must be combined. The new model must accommodate the ancient concern with society's global harmony and the modern concern with personal self-realization.

This new attempt at reconciliation is notoriously associated, in the *Phenomenology*, with the French Revolution. Throughout this relatively short section, Hegel's description echoes Rousseau's conception of a purely participatory democracy, grounded in the abolishment of all forms of political representation, on the elimination of the state's intermediate administrative structures and on an equal distribution of property. In Hegel's terms, Rousseau's ambitious project, as well as the political efforts of Robespierre, Danton or Saint-Just, were aimed at the attainment of 'absolute freedom', i.e. a direct equation between personal freedom and the common good. To this end, 'all social groups or classes, which are the spiritual spheres into which the whole is articulated, are abolished', and the purpose of each individual consciousness becomes 'the general purpose, its language universal law, its work the universal work.'³⁴ This very formulation, however, is already indicative of the project's inconsistency. For if personal ideas and deeds are nothing else than general ideas and deeds, their content is simply the formal demand that they become universal. The absolute freedom of Enlightenment is really only the movement whereby determinacy and particularity, which are bound to re-emerge in the course of social life, are cancelled and raised to universality. Since everything concrete is born out of particularity, and perceived therefore as a subversion of the 'general purpose', all that is left is pure negativity. Society degenerates into a police state, obsessed with detecting and denouncing all traces of individuality. The universal harmony promised at the outset leads once again to disharmony, tyranny and death.

This experience drives Spirit to a new peak of self-alienation, which will prompt a new return to itself. Just as the 'disrupted consciousness' sought refuge in faith and pure insight, so too will absolute freedom 'leave its self-destroying reality and pass over into another land of self-conscious Spirit, where ... freedom has the value of truth.'³⁵ This new 'land' is the stage of Morality, the first in a whole new series of phenomenological transformations. Up until now, Spirit was unable to find the harmony it seeks in the concrete realm of social relations. Henceforth, it will attempt to set up a new system of moral values, in which to ground its thoughts and actions. The richness and complexity of this new string of chapters must be reduced here to a series of general indications.

In this new dialectic, Hegel returns to Kantian morality and expands the criticism sketched in the final stage of the dialectic of Reason. His overall strategy is taken over from the dialectic of absolute freedom: in Rousseau's model, particular and general aims were immediately joined together but not truly synthesized, which led to their final divorce; in Kant's account of practical reason, moral virtue is likewise juxtaposed to the concrete element of human life. Kant's moral doctrine is founded on the antithesis between the actual content of human choices and what that content

³⁴PS, 357 / HW 3, 433.

³⁵PS, 363 / HW 3, 441

ought to be. Since moral goodness is defined as that which all actions must aspire to but can never really attain, it is bound to remain a pure *beyond*. This line of argument is illustrated by a critique of Kant's three postulates of practical reason: the practical need to believe in personal freedom, the practical need to believe in immortality and the practical need to believe in the existence God.³⁶ For Kant, although these beliefs cannot be theoretically demonstrated, they are the necessary foundations of every moral judgement. If human life were confined to a purely eudaemonistic framework, unpleasant choices and decisions, however moral, would never be made. For morality to prevail over happiness, its motivation must lie outside the eudaemonistic sphere, in the ideas of freedom, immortality and God.

First, Hegel argues that the Kantian postulate of freedom arises from the spurious presupposition that the natural world, governed by natural causality, is separated from the moral world, governed by moral mandates. This separation is spurious because reality's agreement with morality is grounded in human actions themselves, and not in some abstract postulate. Although moral consciousness 'starts from the idea that, *for it*, morality and reality do not harmonize, . . . it is not in earnest about this, for in the deed the presence of this harmony becomes *explicit for it*.'³⁷ Consciousness is not 'in earnest' about morality because what it deems morally necessary is ultimately also what it wants to see happen. 'What it really holds to be most desirable, to be the Absolute, is that the highest good be accomplished, and that moral action be superfluous.'³⁸

This 'duplicity' (*Verstellung*) is also brought out in Hegel's critique of the postulate of immortality, which locates moral goodness at an infinite distance, beyond the confines of human existence. Once again, this attitude is grounded in a separation which is not taken seriously. Since the truth of morality is shifted (*verstellt*) to a 'nebulous remoteness, where nothing can any more be distinguished or comprehended', the truth of morality is really only the imperfection of actual existence, 'supposed to be a *progress towards* perfection.'³⁹ Ultimately, then, the moralist 'is concerned not about morality, but solely about happiness as such.'⁴⁰

Finally, regarding the existence of God, Hegel moves away from Kant's actual argument and bases his critique on the distinction between a 'pure duty', corresponding to a universal moral imperative, and a variety of 'determinate duties' conveying what is morally right in each particular situation. Faced with this distinction, morality must find a way of harmonizing the contingency of specific duties (viz. the obligation to help one's friends or relatives) with the absolute moral necessity of pure or universal duties (viz. his obligation to help every human being, or to end human suffering). This opposition is bridged with the aid of a divine consciousness, or a divine lawgiver, to whom the moralist entrusts the role of sanctifying particular

³⁶See Kant's *gesammelte Schriften*, vol. 5, 132.

³⁷PS, 377 / HW 3, 456

³⁸*Ibid.*

³⁹PS, 378 / HW 3, 458

⁴⁰PS, 379 / HW 3, 459

duties and rendering them necessary. For Hegel, however, this new expedient amounts to a new form of duplicity, which opens the door to an eudaemonistic contamination of the whole model. By legitimating, through God, the adaptation of morality to circumstances, the moralist evades the responsibility of discerning the true content of duty and acting accordingly.

A possible way out of these contradictions lies in what Hegel names the standpoint of ‘conscience’ (*das Gewissen*), guided not by abstract mandates, but by specific *convictions*, arisen from concrete challenges and circumstances. Unlike the Kantian model, where the moral sphere is irreducible to particular motivations, this attitude amounts to something like a practical moral judgement, where one’s determinate situation does have a bearing in one’s choices and decisions. In contrast to the moral Spirit, ‘conscience’ observes reality, acts on it, acknowledges the consequences of its actions and draws its own conclusions. But though this kind of pragmatism is more operative than Kant’s solution, it is still vulnerable to the criticism levelled against the entire moral sphere: insofar as conscience’s convictions are its own—i.e. personal and ultimately arbitrary—they are bound to clash with the convictions of other consciences, and are hence unable to provide the basis for a ‘universal law of action’.

In the dialectic of virtue, the opposition between individual beliefs and the ‘way of the world’ led to a discussion of monarchy, despotism and revolution. This time, the same conflict is discussed in light of the individualism characteristic of the Romantic period. On one side, conscience is immediately sure (*gewiss*) of its insights. Its self-confidence evolves into a special regard for its own judgements and intuitions—a ‘moral genius’ which it longs to see recognized.⁴¹ However, precisely because these insights are mere intuitions, their truth cannot be fully shared. When made public, they are met with distrust and incomprehension. In its efforts to overcome this difficulty, conscience refrains from acting, so as to avoid falling into contradiction or betraying its inner convictions. Retreating into itself, it ‘lives in dread of besmirching the splendour of its inner being by action and existence; and, in order to preserve the pureness of its heart, it flees from contact with the actual world, and . . . vanishes like a shapeless vapour that dissolves into thin air.’⁴² In doing so, however, it ends up becoming the opposite of what it first appeared to be: instead of a concrete ethical Spirit, conscience becomes as vague and elusive as the moralism it first criticized.

This solitary conscience—this ‘beautiful soul’, bent on remaining true to its beliefs—is confronted with the anonymous mass of society—a ‘universal other’ bound to the established moral order.⁴³ While the former refuses to abide by social conventions, which it sees as falling below its inner standards, the latter condemns

⁴¹ PS, 397 / HW 3, 481

⁴² PS, 400 / HW 3, 483f.

⁴³ The term *schöne Seele* is borrowed from Goethe’s *Wilhelm Meisters Lehrjahre*, which together with Rousseau’s *Confessions* seems to be on the back of Hegel’s mind throughout this section. See Norton, *The Beautiful Soul*; or Paha, *Die schöne Seele Hegels und die Literatur der Frühromantik*.

the former's self-isolation and resents its sanctimonious attitude. For the conscientious individual, the others are but critics who fail to grasp his or her own integrity. For the crowd, conscience is *sinful* and must be brought back from isolation. The two sides are pitted against each other and can only come together through a new and more sophisticated form of recognition, which leads to the conclusion of the dialectic of Spirit. In this final and difficult section, Hegel speaks of reconciliation in terms of *confession* and *forgiveness*. In the course of the argument, both parties are shown to be essential to the definition of Spirit, but incapable, on their own, to provide its truth. Whereas the individual 'I' of conscience stands for pure self-certainty, devoid of actuality, the universal 'I' of society stands for pure actuality, devoid of self-certainty; whereas the first moment is Spirit *for itself*, the second moment is Spirit *in itself*. Their final reunification comes about when conscience *confesses* its obstinacy to the other, and becomes part of it, and when the other *forgives* this obstinacy and embraces conscience. The result is described as follows:

The reconciling *Yes*, in which the two 'I's let go their antithetical *existence*, is the *existence* of the 'I' which has expanded into a duality, and therein remains identical with itself, and in its complete alienation and opposite possesses the certainty of itself: it is God manifesting in the midst of those who know themselves as pure knowledge.⁴⁴

The sinuosity of this final section, which can be interpreted in different ways, should not make us lose sight of the dialectic's global meaning. Throughout the chapter devoted to Morality, Spirit claimed to be 'certain of itself' in different, but equally inconsistent ways: it looked for its essence in a series of moral postulates, in the immediacy of concrete existence and in the purity of its own convictions. In all of these cases, the self remained self-divided. Finally, cut off from the objective world and from other subjects, it was led to realize that self-certainty is but one moment of Spirit. To be 'certain of itself', Spirit must also be certain of its world. Or rather, it must find itself in its world as both subject and object, individual and universal, certainty and truth. Only the acknowledgement of this reciprocity can reveal Spirit's true identity: 'the human subject who creates his own history, who no longer finds the universal outside of himself, who no longer opposes the universe, but carries it within and absorbs it into himself.'⁴⁵

As Hegel puts it, Spirit's individual and universal moments have said 'yes' to one another. But he also suggests that this unity, albeit complete, was accomplished at a horizontal level, and not yet at a vertical one. The unity of the self with its world has assumed the form of a living social organism, but its life is still a worldly or secular life. It has yet to work out its relationship with divinity, left unresolved in the previous stages of the progression. This is why Spirit's 'pure knowledge' finds 'God manifesting itself in its midst', and why this new standpoint still calls for a higher form of synthesis.

⁴⁴PS, 409 / HW 3, 494

⁴⁵Hyppolite, *Structure et Génèse de la Phénoménologie de l'Esprit*, 477

8.3 Tragedy and Comedy

The relationship between reason and faith was introduced with the unhappy consciousness, discussed in connection with the ethical world and redefined with the Enlightenment's critique of religion. Since all of these stages led to contradiction, Spirit must now revive this debate and look for a higher form of reconciliation. This is the task undertaken in the stage of Religion, the last in the progression.

The dialectic begins with the simplest form of religious experience, associated with the natural world. The contact between the human and the divine realms is initially mediated by sunlight, the most primitive expression of divinity. Afterwards, this contact evolves into the worship of plants and animals. Finally, it leads to the fabrication and worship of man-made artefacts, whose initial geometric shape is gradually replaced by the more concrete depiction of animals and human figures. But although these objects are less abstract than the previous ones, they still offer a limited image of divinity. They represent god as a mere 'other', opposed to the worshipper and confined to a chunk of wood, a piece of clay or a block of stone. For a higher communion between the human being (Spirit *for itself*) and god (Spirit *in itself*), artisans must put themselves into their work and transform the object into a living expression of self-consciousness. Only by being 'purged of the immediate shapes of Nature' and 'dissolved into spiritual shape'⁴⁶ can the object become a true meeting point of the self's earthly existence and its divine calling: the concrete interface 'where Spirit meets Spirit', or 'the unity of self-conscious Spirit with itself.'⁴⁷

In response to this demand, the artisan becomes an *artist* and 'natural religion' is replaced by 'religion in the form of art' (*Kunstreligion*). Spirit will now assess different forms of artistic expression and the extent to which they succeed in bridging the gap between the human and the divine. Starting with the adoration of sculpted deities and the intonation of religious hymns, it moves on to the synthesis of image and speech provided by religious cults and festivals. Finally, it embraces the properly spiritual experience afforded by the literary work of art, where gods and men are put in direct relation with each other. This dialectical crescendo also reflects the development of an increasingly social religious model: whereas natural religion still allows for a private exchange between one believer and his or her god, religious cults and festivals are shared by a whole nation. Literary works, on the other hand, are neither personal nor national, but universal: their raw material is the 'single pantheon' of human language and their main subject humanity as such.⁴⁸

The three literary genres examined by Hegel—epic, tragic and comic poetry—illustrate different conceptions of the relationship between self-consciousness and a divine or transcendent being. More than mere descriptions of specific art forms, these three models represent three different answers to the questions initially raised in the

⁴⁶PS, 424 / HW 3, 512

⁴⁷PS, 424 / HW 3, 511f.

⁴⁸PS, 439 / HW 3, 529

stage of unhappy consciousness and left open ever since: How free or independent is self-consciousness? Can the self really rise above all forms of alienation and become purely *for itself*? Is it bound to relinquish its independence and submit to an absolute *in itself*? Or is it doomed to remain trapped in this contradiction, thrust back and forth from self-unity to self-division?

In epic poetry, this dilemma takes the form of a syllogism ‘in which the extreme of universality, the world of the gods, is linked with individuality, with the bard, through the middle term of particularity.’⁴⁹ This middle term is filled by the deeds of epic heroes, whose concrete individuality, raised to an ideal form, embodies the values of an entire people. In the course of the narrative, however, this heroism is shown to be entirely dependent on the will of the gods. Epic heroes may see themselves as free, self-determining individuals, but the epic tale invariably proves them wrong. The gods are the real source of determination, the ‘substantial powers’ that rule human life. But this is only one side of the story. Since the gods themselves are conceived as individuals, animated by the same kind of passions and desires that animate human beings, their power is not as substantial after all. In a ‘comical self-forgetfulness of their eternal nature’, they also feel love, rage, envy and impotence. In short, they are also *human*, and the hero’s confrontation with divinity becomes, therefore, ‘an empty and arbitrary showing-off’, which ‘transforms the apparent earnestness of the action into a harmless, self-confident play.’⁵⁰

Following this realization, the syllogism’s middle term is eliminated and its extremes are redefined. As regards universality, the authority of the gods is transferred to a higher sphere, to which the gods themselves are subordinated, namely the law of fate. This is now the divine moment of the relationship, the ‘substantial power’ facing the individual. As regards individuality, the self-confidence of the epic hero is transferred to the bard, who is led to abandon his passive role and to take part in the action. Unlike the epic hero, he ‘is himself the speaker, and the performance displays to the audience . . . *self-conscious* human beings who *know* their rights and purposes . . . and know how to *assert* them.’⁵¹ The epic dispute becomes a *tragic* dispute, thereby reviving the opposition between human and divine laws discussed earlier. This time, however, the stakes are much higher. Whereas the antagonistic stances illustrated by Creon and Antigone stood for immediate ethical mandates, the tragic standpoint that will now be considered is a concrete spiritual standpoint, which has tested, negated and reabsorbed all of the *Phenomenology*’s previous developments. This point is extremely important. The kind of knowledge that will now be put to the test is not simply the knowledge held by ordinary consciousness, or the knowledge concerning a specific situation or course of events. It is the necessary culmination of the long and complex journey we have hitherto discussed. Consciousness, Self-consciousness, Reason, Spirit, Religion, Art: what is now at stake is *all of this*. The confrontation between the tragic hero and his or her fate is a test to human knowledge as a whole, in all its dimensions.

⁴⁹PS, 441 / HW 3, 531. Translation modified.

⁵⁰PS, 442 / HW 3, 533. Translation modified.

⁵¹PS, 444 / HW 3, 534

In the beginning of this book, when introducing the relationship between freedom and knowledge, I evoked Aristotle's conception of tragedy, and particularly the notion of *περιπέτεια*, i.e. the idea of a complete *reversibility* of human knowledge. This same idea is at the centre of the present stage of the *Phenomenology*, where the essence of tragedy is defined precisely as the 'inversion of the known into its opposite', or the 'changing round of the rightness of character and of knowing into the rightness of the opposite.'⁵² For Hegel, as for Aristotle, what defines the tragic experience is the movement whereby Spirit, convinced it knows the truth about itself and the world, is disabused of this conviction. The disillusionment experienced countless times before is repeated once again, but this time at the very summit of the progression. As the tragic situation unfolds, Spirit is again caught up in the antithesis, which it claimed to have left behind, of certainty and truth. What it took to be true is not true, and what it took to be untrue is true after all.

This *περιπέτεια* (*Verkehrung*, *Umschlagen*) is not a regional or contingent phenomenon, which could be corrected through a mere change of content. This impossibility is immediately made clear by Hegel's description of the other side of the equation. While the hero's knowledge is contradicted by fate, the latter is just as one-sided, and could just as easily have been different. In other words, if the hero could have anticipated the actual result of his or her actions and altered their content, fate would still have had the final word, for it is *by definition* opposed to the individual. What is learned through tragedy is not simply that the knowledge held by Spirit is wrong, but that it is *intrinsically open to the possibility of error*, and *intrinsically unable to eliminate this possibility*. In tragedy, '[consciousness] pays the penalty of trusting a knowledge whose ambiguity, *for such is its nature*, also becomes explicit *for* consciousness and a *warning* to it.'⁵³ The hero learns that *all knowledge*, be it individual or universal, subjective or objective, earthly or divine, is equally one-sided and liable to change to its opposite.

The truth . . . of the opposing powers of the content and of consciousness is the result that both are equally right, and therefore, in their antithesis, which is brought about by action, both are equally wrong. The course of the action demonstrates their unity in the natural downfall of both powers. . . . Both are *oblivion*, the disappearance of actuality and the action of the powers of substance, . . . for none is for itself the essence.⁵⁴

The experience of tragedy is hereby consummated. The stage is clear, as it were, and consciousness is left alone with itself. It has seen everything solid or substantial dissolve before its eyes, every notion and every law lose their value. Truth, divinity and fate have all been exposed as the props of a fictional world, imagined and sustained by self-consciousness, and all that is left for the hero is to 'step forth from his mask', acknowledge the deception *as deception* and exhibit *himself* as 'the fate of the gods . . . and of the absolute powers themselves.'⁵⁵ This global recognition

⁵²PS, 447 / HW 3, 538

⁵³PS, 448 / HW 3, 539

⁵⁴PS, 448 / HW 3, 539f. Translation modified.

⁵⁵PS, 450 / HW 3, 541

(ἀναγνώρισις) marks the transition to the *Phenomenology's* last and most radical form of self-alienation, the final abyss into which all the previous stages are thrown. While tragedy is the *negative* acknowledgement that nothing is what it seems, comedy is the *positive* embracement of this new spiritual order, whose content is precisely the fact that it has none. While tragedy is the emptying out of what was once full and stable, the perplexity of seeing one's world disappear, comedy is the experience of emptiness *as such*, the certainty of this absolute loss. More than that, it is the certainty that there never was anything to be found, and that the progression's previous tribulations were really 'much ado about nothing'.

In the stage of comedy, the seriousness of Spirit's earlier attempts to find itself is turned to ridicule. Consciousness has left the tragic stage: it is no longer engaged in the plot, no longer willing to suspend disbelief. Its standpoint is no longer that of the tragic hero, but that of the actor who has given up his or her role, or the spectator who knows that the play is just a play and that nothing is *really* worth fretting about. What it now sees is a series of *unsubstantial* powers, values and notions propped up as something substantial—and herein lies the source of their comic effect: whatever is deemed sacred, solemn or serious is revealed as a mere *impersonation* of solemnity or seriousness, whose ridicule is directly proportional to its alleged degree of importance.

Although comedy brings out the underlying emptiness of every aspect of human cognition, its derisive power is primarily directed at the 'final' or 'sacred' truths of natural and social life. In the first case, it turns against the purported independence of the empirical world, of its laws and principles. In the second case, it targets the customs and values that shape human society. This double movement is 'the consciousness of the dialectic contained in these maxims and laws themselves, and, consequently, the consciousness of the vanishing of the absolute validity previously attached to them.' As a result, the notions of truth, beauty and goodness 'thus display a comic spectacle: through their liberation from the opinion which contains both their specific determinateness as content and also their absolute determinateness, . . . they become empty, and just for that reason the sport of mere opinion and the caprice of any chance individuality.'⁵⁶

Hegel's conception of tragedy and comedy brings us back, once again, to Plato, the critical tradition and the relationship between freedom and philosophy. And it is not by chance that the figure of Socrates plays such an important role at this stage of the progression. As seen earlier, Plato's dialogues are also built around the notion of *περιπέτεια*, extended to every aspect of human life. As Socrates exposes the 'emptiness' of human opinions, he is describing human existence as a *tragic* existence, in the Hegelian sense of the word. Moreover, since what characterizes human opinions, according to Plato, is not simply the fact they are 'empty', but also the fact that they are usually taken to be 'full'—serious, solemn, substantial—the experience staged in the dialogues is also a *comic* experience, in the Hegelian sense of the word. In the stage of comedy, as in Socrates' enquiries, human ideas are

⁵⁶PS, 451f. / HW 3, 543f.

stripped of their usual ‘mask’ and exposed as the *impersonation* of a series of contents they are unable to convey. But although in the dialogues this experience is just as much tragic as it is comic, for there is ultimately no way of knowing whether reason’s masks do conceal a true face—and, if so, what that face looks like—in the stage of comedy this experience is not as ambiguous. ‘Unlike what happens with tragedy, comedy’s unmasked masks are not the masks of an *αὐτό* [an ‘in itself’] whose density is equal to, or higher than, the density of the masks themselves; they are not the masks of an *enigmatic self-essentiality*, *still left to discover*, which holds everything in its power. No. . . . The mask contains only the mere pretence to be more than it is, and conceals nothing other than the *inanity of this pretence*. . . . [It] only conceals its own emptiness, and is totally in the hands of the subject, who understands it as such and can dispose of it completely.’⁵⁷

The phenomenological experience of comedy is no longer comparable to the perplexity felt by Socrates’ interlocutors throughout his interrogations, nor to the apparent hope displayed by Socrates himself in his ceaseless search for a true conceptual content. The comic self is an utterly disenchanted mode of consciousness, which no longer expects to find, in itself or in others, anything other than emptiness and deceit. Although it is explicitly associated with ‘irony’, its ‘wantonness’ (*Mutwille*) and ‘levity’ (*Leichtsinn*) are more Aristophanic than Platonic, while its ‘contempt’ (*Spott*) and ‘bitterness’ (*Bitterkeit*) vis-à-vis the natural and social realms anticipate Nietzsche’s attitude towards science and morality. What is more, instead of fleeing the acknowledgement of its own condition, the comic self revels in it, as the anarchist revels in anomie, or the nihilist in meaninglessness. Its negative attitude resembles the absolute ‘vertigo’ later described by Baudelaire:

Once the vertigo has entered, it circulates through the air; one breathes the vertigo; it fills the lungs and renews the blood in the ventricle.

What is this vertigo? It’s the absolute comic.⁵⁸

⁵⁷ Carvalho, ‘Verse actuar a sí mismo’, 213

⁵⁸ Baudelaire, *Oeuvres Complètes*, vol. 2, 383

Chapter 9

Beyond the Phenomenology



Abstract Widening the scope of the analysis, I consider the *Phenomenology of Spirit* as a whole and its position within Hegel's system, in order to evaluate the success of its liberating task. First, in Sect. 9.1, I examine the *Phenomenology's* closing chapter, discuss different interpretations of its final outcome and criticize a few misconceptions still associated with Hegel's 'absolute knowing'. Afterwards, in Sect. 9.2, I focus on the work in its entirety and confront the kind of deliverance it provides with the ones aimed at in higher stages of the system, so as to distinguish the different forms and degrees of education and liberation proposed by Hegel. Finally, in Sect. 9.3, after considering different objections levelled against Hegel's project by later authors, I outline Nietzsche's criticisms and his own views on the possibility of a philosophical re-education/release of human consciousness.

Keywords Hegel · Nietzsche · *Phenomenology of Spirit* · Freedom · System · Science · Liberation

The stage of comedy, more than all others, poses a continuity problem. As a result of the dialectic of 'religion in the form of art', Spirit evolved into a purely negative power, whose essence is precisely that it has none and allows for none to be posited. The self has been driven to the edge of itself, condemned to absolute derision. Since its negative power extends to every dimension of human life, any 'serious' stride forward, aimed at a new and less sterile attitude, is bound to be greeted with laughter and contempt. As Hegel puts it, 'what this self-consciousness beholds is that whatever assumes the form of essentiality over against it, is instead dissolved in it ... and is at its mercy.'¹

If the progression were to end here, the truth promised at the outset would remain out of reach. The *Phenomenology's* painstaking survey would lead to a logical and existential cul-de-sac and its liberating task would prove impossible. Freedom would be reduced to the negative freedom of comedy and Hegel's project to yet another

¹PS, 452 / HW 3, 544

critical project, culminating in a peculiar form of scepticism. This, however, is not Hegel's position: the standpoint of comedy is only at first perceived as a dead end. As so often before, the self's complete self-alienation will be revealed, on closer inspection, as the moment of a wider dialectical synthesis, in which selfhood and otherness are once again reunited.

This final reconciliation takes place in the stage of 'revealed religion', right before the *Phenomenology's* grand finale. Throughout a dense and demanding succession of paragraphs, Hegel makes two important claims: first, he argues that the way out of comedy's absolute derision is absolute *self*-derision, which leads to the discovery of a new form of unhappy consciousness and a new mode of religious experience; second, he claims that the latter is really only a replacement for, or an anticipation of, Spirit's true form. The Christian notion of a Holy Trinity, in which God is simultaneously human and divine, individual and universal, prompts the realization on the part of consciousness that it is *itself* the meeting point of these conflicting moments. The contradictions that have plagued the previous stages are provisionally resolved by Christian theology and sublated at last by a truly dialectical mode of cognition, boldly entitled 'absolute knowing'.

In the progression's final stage, the phenomenological subject abandons the doxastic realm and steps into a new cognitive sphere. The 'exposition of the untrue consciousness in its untruth' is finally complete and what is now left to discover is simply the negative side of untruth, i.e. science or actual knowledge. But despite Hegel's clear architectonic scheme, this conclusion raises a number of difficult questions. What exactly is 'absolute knowing'? Or rather, what does it mean to have reached such a standpoint? If consciousness has indeed traversed the entire realm of appearance and left behind each and every form of self-deception, what does its new outlook actually amount to? These questions are vital to understanding the meaning and scope of Hegel's project, and not only from an epistemological perspective. Given the correlation between the possibility of knowledge and the possibility of a free or empowered life, which has been the guiding thread of this enquiry, we must also ask what this conclusion means with regard to freedom. Has consciousness, in reaching the stage of 'absolute knowing', secured the cognitive release envisaged at the beginning of the progression? And, if so, what does its newly found freedom amount to?

In the following sections, I attempt to shed light on the *Phenomenology's* answers to these questions. As will become clear, however, they cannot be fully clarified without considering a further and no less complex aspect of Hegel's enterprise, namely the position occupied by the *Phenomenology* within his wider philosophical system. On the one hand, the phenomenological progression is conceived and described as the elevation towards science and truth; on the other hand, it is characterized by Hegel as a mere introduction to philosophy proper, where science and truth are said to develop. This second claim puts the problem of truth and freedom, as discussed in the previous chapters, into a new perspective, whose main implications will be considered below.

9.1 Final Revelations

The Religion chapter can be read as a re-enactment of the itinerary followed by the unhappy consciousness. Here, however, we have already moved past the private relationship between humanity and divinity, as well as the abstract forms of worship, asceticism and self-sacrifice documented earlier. What is described in the dialectic of Religion is the inner division of the concrete Spirit unveiled in the stage of Morality—a shape of Spirit in which individuality and universality have been reunited, but which is still confined to a ‘mundane existence’, and still burdened with a divine calling. Accordingly, the religious conceptions discussed throughout this chapter do not stand merely for individual standpoints, but for global developmental stages, where the phenomenological subject is simultaneously an independent, self-determining agent and the embodiment of a specific cultural context.

Initially, the divine element was located beyond the natural world. God was the essential moment of the relationship, worshipped and summoned by the faithful. Afterwards, through tragedy and comedy, Heaven was gradually ‘depopulated’ and the self emerged as the essential moment of the relationship. The comic consciousness dissolved every external form of authority and became itself the Absolute Being, or the new God. But Hegel will now argue that this new independence is also illusory and destined to implode. Since comedy’s power is pervasive, its poison is bound to contaminate its own source and dissolve the very will that keeps it alive. Indeed, although comedy is ‘the consciousness of the loss of all essential being in this *certainty of itself*’, it is also ‘the loss . . . of this knowledge about itself—the loss of substance as well as the loss of Self.’² In other words, comedy’s derision leads to self-derision, which leads in turn to self-destruction. What was at first perceived as a vast joke gives way to a spiritual void, and the self’s previous arrogance is replaced by misery and humility.

Comedy leads to a new version of the unhappy consciousness, or to ‘the grief which expresses itself in the hard saying that “God is dead.”’³ The ensuing dilemma is by now familiar: Spirit has turned its back on God and religion, but failed to find in itself a worthy replacement. Since the subjective and the objective poles have both proven untenable, they must now be approached together rather than separately. In order to overcome the spiritual void it has encountered, Spirit must cease to regard the human and the divine elements as mutually exclusive. And just like before, the solution lies in the notion of *incarnation*, wherein the two previous conceptions of God—namely God as *absolute Being*, or the abstract deity of natural religion, and God as *absolute Subject*, or the negative power of Comedy—are finally brought together. Unlike these two alternatives, an *incarnate* or *embodied* divinity is simultaneously negative and positive, self-conscious and actual, for-itself and in-itself. Its synthetical nature mirrors the synthetical nature of the human self and replicates its own self-division.

²PS, 455 / HW 3, 547

³Ibid.

With the notion of incarnation, the divine Being is finally *revealed*. The phenomenology of religion, which hitherto sought to accommodate consciousness' self-certainty with the certainty of an absolute essence, placed either outside the natural world or inside the thinking self, has finally given up its dualism. Once God is 'made Man', the unity between humanity and divinity is no longer *strived for* (as in natural religion) or *asserted* (as in comedy), but *directly beheld* (*angeschaut*). In the figure of Christ, consciousness is immediately aware that the divine element is *not* different from the human element—that it was only *posited* as such by the previous forms of experience.

This new development is completely misunderstood if the embodied divinity is regarded merely as an objective presence *endowed* with a divine meaning. If that were the case, this divine individual would only differ from the icons and statues worshipped in the previous stages by its ability to move, think, speak, etc. It would be a living image of divinity, but an image all the same, whose essence would still be located elsewhere, beyond the natural world. Hegel's point, however, is that God is now revealed as a *self-conscious* being—a being that is both conscious of itself as divinity and conscious of its actual existence. In other words, God is now revealed as *Spirit* proper, for Spirit is precisely 'the knowledge of oneself in the externalization of oneself; the being that is the movement of retaining its self-identity in its otherness'⁴ This declaration is the turning point of the whole dialectic: Hegel is now claiming that consciousness, in having arrived at a truly dialectical conception of divinity, *has simultaneously arrived at a true conception of itself*. The contradiction between self-certainty and transcendence it was unable to resolve has now been sublated by a new religious shape, which emerges before consciousness as in a mirror and reveals at last its true identity. God is immediately *present* as Spirit,⁵ and through God's presence consciousness will also be revealed as Spirit—no longer the incomplete Spirit that had come out of the Morality chapter, but a Spirit that has fulfilled its transcendent calling and joined together the human and the divine realms.

Once the religious consciousness realizes that the incarnate divinity is identical to itself, Spirit is at last recognized as Spirit and the progression attains its final stage. For now, however, we still have the human self opposed to the divine self, not yet aware of their coincidence. For the former to see itself in the latter, or for the latter to come out of the mirror, as it were, and join the former, two last steps are needed. First, the image in the mirror is not yet fully faithful to the original because the embodied divinity is still a particular individual. As we have seen, however, the religious consciousness is no longer simply an individual standpoint, but the individual reflection of a universal standpoint. In order to correct this discrepancy, God must be divested of His individuality and raised to a universal form. He must die and be resurrected, abandon His human body and re-emerge as the 'Holy Spirit'. In this new form, God's presence is not eliminated, but expanded: the divine Being reappears before consciousness as a truly universal Self, whose presence extends

⁴PS, 459 / HW 3, 552

⁵PS, 461 / HW 3, 554

to the whole world. His death ‘becomes transfigured from its immediate meaning, viz. the non-being of this *particular* individual, into the *universality* of the Spirit who dwells in His community, dies in it every day, and is daily resurrected.’⁶

The image in the mirror is now fully formed, but Spirit’s return to itself is still dependent on one final task. According to Hegel, the language employed by religion is still anchored in representational thinking (*vorstellendes Denken*), and not yet in notional thinking (*begreifendes Denken*). This distinction brings us back, yet again, to the methodological distinction of the preface: although religion has unveiled the true nature of Spirit, it has done so by means of different representational predicates (*Vorstellungen*), corresponding to different spiritual events. The content of religion is ‘the true content, but all its moments ... have the character of being uncomprehended, of appearing as completely independent sides which are externally connected to each other.’⁷ In short, religion still views Spirit in a *narrative* way, as a succession of transformations predicated of the same divine Being. This kind of perspective, however, is not yet Spirit’s native perspective. It is still a borrowed perspective, which has yet to grasp the *immanent* and *simultaneous* nature of these transformations. Although religion has helped Spirit rise to its true form, its phenomenological role has come to an end. Like all of the progression’s previous stages, its specific language has outgrown its use and must now be replaced by a new mode of expression.

Since Spirit was brought face to face with its own image, all that it must do now is step forward and reclaim its authorship. With this final movement, what in religion was still ‘*content*, or a form for representing an other’, is revealed as ‘the *Self*’s own *act*.’⁸ Humanity and divinity, individuality and universality, subject and substance—all of these contrasts are divested of their narrative form and disclosed as the interchanging moments of the same overarching Notion. They are no longer merely placed before consciousness, but absorbed by it and dissolved in its inner movement.

Later, in the *Encyclopaedia*, Hegel will divide the last stage of the progression, entitled ‘Absolute Spirit’, into the substages of Art, Revealed Religion and Philosophy. Here, the division is not as straightforward, but the main argument is largely similar. The standpoint of religious revelation, echoed by Christianity’s Trinitarian theology, is the one that best articulates the main dialectical contrasts encountered throughout the *Phenomenology*. By positing a divine or absolute Being (God, the Father) that is both human (Christ, the Son) and universal (the Holy Spirit), the Christian worldview manages to reintegrate, reconceptualize and overcome all the one-sided solutions that were tested and abandoned in the previous stages. The dialectical form of identity Hegel terms Spirit, introduced and discussed ever since the self-consciousness chapter, is hereby fully formed.

⁶PS, 475 / HW 3, 571

⁷PS, 463 / HW 3, 556

⁸PS, 485 / HW 3, 582

According to Michael Forster, Hegel's 'whole system is conceived, on one level at least, as a defence or rational reworking of the Christian conception of God.'⁹ And the privileged position accorded in the *Phenomenology* to 'revealed religion' seems to corroborate this idea. But the dialectic of religion is not the final rung in the progression—just as the stage of Religion is not the conclusion of the *Encyclopaedia*. And if the current stage is the progression's most 'pious' moment, the next one is in a way the most 'impious'. After hailing the religious worldview (or the phenomenological equivalent of the Christian worldview) as the most sophisticated of all, Hegel goes on to claim that it is still a mere crutch, so to speak, used by consciousness for support, or a mere prelude to the acknowledgement of its true nature. As soon as consciousness realizes that it is itself all of the things it ascribes to the religious Spirit—namely subjective *and* objective, individual *and* universal, human *and* divine—it has no further need for the religious notions it came to rely on. In other words, Hegel sees the Christian standpoint as a mere preparation for his own philosophical standpoint. It is not the *Phenomenology* that amounts to a philosophical translation of Christianity, but the other way around: the Christian standpoint is itself a mere translation of philosophy, destined to be sublated and replaced by the real thing.

Only in the progression's final standpoint is the Notion pursued by consciousness finally *realized* (*realisiert*). While in the previous stages knowledge and truth kept falling outside of each other, leading to new and increasingly complex cognitive models, their separation is now overcome. This coincidence, however, is not a simple or static coincidence, but a *dynamic* or *infinite* one. The truth revealed at the end of the progression is no longer a true content or object placed before consciousness. This kind of interpretation leads back to the ingrained dualism that has now been left behind. The truth unveiled in this final stage consists rather in the emergence of a truly systematic mode of cognition, in which all previous dichotomies have been eliminated, and where everything is joined together with everything else. On the other hand, neither does this mean that difference and variety have been eliminated, or that experience was reduced to a simple reiteration of the same universal content. Knowledge and truth have been united, but this unity is a *self-moving* unit, in which identity is achieved through a continual return to otherness. Spirit realizes at last that the movement whereby it places an other before itself—be it an object, or another individual, or the whole world, or a divine being—is not simply a *means* or a *vehicle* for the attainment of self-recognition. This movement is *itself* the truth of Spirit, the absolute mediation that is its very essence. Spirit 'is in truth actual only in so far as it is the movement of positing itself, or is the mediation of its self-othering with itself. . . . It is the doubling which sets up opposition, and then again the negation of this indifferent alterity and of this antithesis. Only this *self-restoring* sameness, or this reflection in otherness within itself . . . is the True.'¹⁰

⁹Forster, *Hegels Idea of a Phenomenology of Spirit*, 12

¹⁰PS, 10 / HW 3, 23

Absolute knowing is reached when every remaining *Vorstellung* is dissolved into the absolute flow of *Begriff*. And this movement encompasses all the main dichotomies we have hitherto considered: with regard to the contrast between subjectivity and objectivity, Spirit has become its own object, or ‘the movement of retaining its self-identity in its otherness’¹¹; with regard to the contrast between subjectivity and intersubjectivity, Spirit has become both ‘an individual Self’ and the ‘universal Self, or the Self of everyone’¹²; with regard to the contrast between humanity and divinity, Spirit has emerged as the truth of religion, or ‘the Divine’s intuition of itself.’¹³ In its final form, Spirit is the reconciliation of these three main syntheses, joined together in the element of science.

9.2 Science and the System

With ‘absolute knowing’, consciousness has reached the stage of science. But what exactly does this mean? And how does it enable the cognitive release originally envisaged by Plato and discussed in the previous chapters? To start answering these questions, it is useful to focus on the word science, or *Wissenschaft*, and the different meanings it assumes throughout the *Phenomenology*. In the introduction, we are told that science is not only the result of Hegel’s progression, but also the process leading up to it: ‘the way to Science is itself already Science, and hence . . . the Science of the *experience of consciousness*.’¹⁴ What Hegel means by this, as we have seen, is that the *Phenomenology* is not a random or contingent progression, but a necessary one. It consists in the self-realization of the same overarching Notion, whose final stage is, and can only be, ‘absolute knowing’. But apart from being a process, Science is also a result. And sure enough, in the following paragraph of the introduction, Hegel makes clear that only in the progression’s finishing stage, ‘where appearance becomes identical with essence’, does its exposition coincide with ‘the authentic science of Spirit.’¹⁵

Let us now focus on this second and more robust definition of science. Like Plato’s ἐπιστήμη, science as a result is the opposite of opinion, or δοξάζειν. It corresponds to a standpoint that no longer merely claims its understanding of reality to be true, but that *knows* it to be so. And the attainment of this knowledge can only be achieved in a negative, or dialectical manner: in the *Phenomenology*, as in Plato’s dialogues, truth can only emerge, if at all, once consciousness has exposed and discarded all the ‘false truths’ it is burdened with. Therefore, when Hegel claims that

¹¹ PS, 459 / HW 3, 552

¹² PS, 462 / HW 3, 555

¹³ PS, 483 / HW 3, 580

¹⁴ PS, 56 / HW 3, 80

¹⁵ PS, 57 / HW 3, 81

consciousness, at the end of the *Phenomenology*, has achieved the standpoint of science, he is also claiming that the 'natural ideas, thoughts and opinions' with which consciousness is initially 'hampered' have all been removed, and there is no longer any illusion at work in human life.

Given the boldness of this statement, it is important to recall, if only very briefly, the long series of lower-order requirements that need to be fulfilled for it to be true. First, the elevation to the standpoint of science, in the terms we have just considered, is conditional on the correct identification of every form of untruth subscribed to, whether implicitly or explicitly, by human consciousness. If Hegel's final claim is justified, the vast kaleidoscope revealed throughout the *Phenomenology* must have succeeded in covering not only every cognitive angle or attitude available throughout human life, but also every combination of angles or attitudes available to us. If this inventory is not complete, there might still remain within us a hidden presuppositive core, incompatible with Hegel's final purpose. Second, for Hegel to be right, he must have correctly identified the dialectical sequence whereby all of these different angles and attitudes follow and replace one another. If this chain was incorrectly assembled, the progression's internal necessity is liable to be compromised, along with its final goal. Third, the elevation to science also requires that all of the *Phenomenology*'s stages have been thoroughly sublated and overcome, so that no presupposition was allowed to exceed its proper life-span. As we have seen when discussing Hegel's methodology, if each new sublation does not correspond to a whole new worldview, its critical role is irrevocably compromised. Fourth, in line with the previous requirements, if consciousness has indeed been raised to the stage of science it must be able to guarantee, beyond a shadow of a doubt, that the final phenomenological stage is indeed the last one, and not the product of a hidden set of assumptions, which might extend the progression into new and unknown territories. Fifth, although all of the aforementioned conditions refer to the progression itself, and to how it is crafted, they must also be applied to the reader or interpreter. For even if the *Phenomenology* itself does fulfil all of the previous requirements, we readers may not be able to understand or assimilate all of its elements and arguments. In this case, even though we may be provided with all the means necessary to arrive, with Hegel, at the standpoint of science, we may be incapable of doing so, out of ignorance or inattention.

Furthermore, this long list of requirements is decisive not only for the attainment of truth, but also for the attainment of freedom. The cognitive liberation envisaged by Hegel also hinges on whether his work succeeds in identifying each and every form of self-deception, on the ability to retrace the dialectical sequence that unites all forms of self-deception, on the power to criticize and sublate all of these forms, on whether the phenomenological progression is in fact a finite progression and on the actual ability to comprehend and assimilate all of its arguments and conclusions. Again, if any of these demands is left unanswered, the freedom promised at the outset will remain out of reach. For even if consciousness manages to overcome some, or even many of the obstacles along the way, it might still be caught in a sophisticated form of self-deception, and thereby prevented from thinking and acting freely.

But this is still not all. The assessment of the *Phenomenology*'s final result also requires that one look beyond the work itself, at Hegel's wider philosophical system. For although the *Phenomenology* is defined as the educational path towards science and truth, Hegel also claims that the progression is a prelude to science proper, i.e. to the actual development of truth. In a much-discussed 1807 letter to Schelling, referring to the *Phenomenology*, Hegel declares that he has not yet gotten 'beyond the introductory stage into the heart of the matter.'¹⁶ And the same idea is reiterated in the work itself—both in the final paragraphs, where Hegel discusses what has been achieved and what is left to do, and in the preface, written afterwards, with the *Logic* already in mind.

Let us focus on these texts. Towards the end of the *Phenomenology*, Hegel insists on the qualitative difference between the realm of untruth, or appearance (*Erscheinung*), still 'burdened with the difference of consciousness [from its object]' and the realm of truth, or science, where the object is no longer 'in itself' but 'its own restless process of sublating itself.'¹⁷ The absolute entwining of self-identity and otherness which Hegel calls the Notion (*Begriff*) has finally been achieved and consciousness has at last been raised to truth. But although this achievement is the progression's final result, it is not yet the end of Spirit's formative path. As Hegel proceeds to argue, the element of science is both an ending and a new beginning—the gateway to a new cognitive element, which has yet to be fully understood and explored. In other words, while the 'exposition of the untrue consciousness in its untruth' is still but a negative procedure, what now emerges is the possibility of a new kind of exposition, consisting in the attainment of *actual* knowledge. Hegel thus introduces yet another definition of science, irreducible to the ones we have previously considered: apart from science as a *process* and science as a *result*, the term *Wissenschaft* now stands for a higher form of cognition. It is also a process, but one that takes place *within* the element of science.

This new, post-phenomenological task is described in the following terms:

Spirit, therefore, having won the Notion, displays its existence and movement in this ether of its life and is Science. In this, the moments of its movement no longer exhibit themselves as specific *shapes of consciousness*, but ... as *specific Notions* and as their organic self-grounded movement.¹⁸

This movement does not belong in the *Phenomenology*, but in the novel and more sophisticated enquiry expounded in Hegel's *Logic*. While in the phenomenological progression knowing and truth are still separated, in science proper knowledge 'is freed from its appearance in consciousness'¹⁹ and revealed as pure self-movement.

This idea is restated in the preface, where the *Phenomenology* is again described as a prelude or a preparation to science. Although its purpose is to raise consciousness to the standpoint of truth, the latter is not yet the end of reason's

¹⁶Hegel et al., *Briefe von und an Hegel*, 161

¹⁷PS, 491 / HW 3, 588. Translation modified.

¹⁸PS, 491 / HW 3, 589

¹⁹*Ibid.* Emphasis added.

self-development. In other words, ‘absolute knowing’ is not yet ‘absolute knowledge’: it is both a result and a new starting point, the transition from a negative to a positive form of philosophizing.

With this [i.e. the unity of knowing and truth], the Phenomenology of Spirit is concluded. What Spirit prepares for itself in it, is the element of [true] knowing. In this element the moments of Spirit now spread themselves out in that *form of simplicity* which knows its object as its own self. They no longer fall apart into the antithesis of being and knowing, but remain in the simple oneness of knowing; they are the True in the form of the True, and their difference is only the difference of content. Their movement, which organizes itself in this element into a whole, is *Logic* or *speculative philosophy*.²⁰

In the introduction to the *Science of Logic*, the direct follow-up to the *Phenomenology*, this connection is made even more explicit. Hegel claims, in similar terms, that the outcome of the *Phenomenology* is the ‘notion of pure science’, but not yet its ‘content’.²¹ Unlike ‘absolute knowing’, the Logic is not merely the discovery of a true or scientific standpoint, but the exposition of ‘pure knowledge in the full compass of its development’.²² Interestingly, Hegel also frames the argument in terms of the *freedom* acquired by means of this transition: whereas ‘pure science thus presupposes the *liberation* from the opposition of consciousness’,²³ it is up to the Logic to acknowledge and actualize this newly found freedom.

This upgrade from notion to content entails a new dialectical progression, with a new starting point and a new ending. In the *Phenomenology*, experience was set in motion by the acknowledgement that the most immediate form of knowledge was in fact the most abstract. Consciousness was separated from its object and confined to an unfree mode of thinking. In the Logic, this opposition has been overcome. The starting point is no longer a specific version of things, or a specific ‘shape of consciousness’, but a ‘pure notion’. Yet like the *Phenomenology*, the Logic follows its own immanent dialectic. Starting from the notion of ‘being’—the simplest and richest of notions, inherent in every single thought—the logical subject stumbles into a new series of contradictions. These do not stem, however, from the disagreement between a specific concept and its objective translation. ‘Logic has nothing to do with a thought *about* something which stands outside by itself as the base of thought.’²⁴ Rather, this movement is fuelled by the notion’s internal contradictions, which are both subjective and objective, ideal and actual.

Finally, it is also worth considering Hegel’s characterization of the transition from the *Phenomenology* to science in the *Encyclopaedia*, the definitive exposition of his philosophical system. In the introduction to the *Encyclopaedia*’s first part, the Logic, the phenomenological progression is once again described as an introductory step, ‘meant to contribute principally to the insight that the questions one entertains and

²⁰PS, 21f. / HW 3, 39

²¹SL, 29 / HW 5, 42. Translation modified.

²²SL, 47 / HW 5, 66

²³SL, 29 / HW 5, 42. Emphasis added.

²⁴SL, 29 / HW 5, 43

holds as utterly concrete in the representation of the nature of *knowing*, *faith* and so forth in fact lead back to *simple* thought-determinations that receive their definitive treatment only in the Logic.²⁵ The role of the phenomenological progression is to liberate consciousness from its usual self-dividedness and to reveal the pure freedom of logic.

In logic, thoughts are considered in such a way that they have no other content than that which belongs to and is generated by thinking itself. In this way, the thoughts are *pure* thoughts. Thus spirit relates purely to itself and is therefore free, for freedom is precisely this: to be at home with oneself in one's other, to be dependent upon oneself, to be the determining factor for oneself.²⁶

Despite Hegel's insistence, at the end of the *Phenomenology*, on the absolute reciprocity of subject and object, this and similar passages seem to indicate a primacy of the former over the latter. Indeed, insofar as the Logic deals with pure, self-sufficient thoughts, independent of factual or contingent considerations, it may well be viewed as a purely subjective exercise. If this were the case, the liberation achieved by the *Phenomenology* would ultimately amount to an elimination of objectivity, in favour of a strictly mental mode of cognition. But this kind of criticism only takes into account the first part of Hegel's argument. Although the Logic is indeed the domain of pure thinking, it is only the opening stage of a new dialectical journey. More precisely, it is the *subjective* moment of this wider journey, to be followed by the *objective* stage of Nature and the *subjective-objective* stage of Spirit.

In other words, what is revealed, at last, in the Logic is the *true* or *scientific* mode of thinking—the mode of thinking which every true thought must follow, regardless of its stage of development. And its first instantiation is the notion of being: i.e. being *as such*, prior to any specific qualification. The Logic starts from the assumption that everything *is*, regardless of being *this* or *that*, and this basic truth is said to underlie all forms of meaning. As usual, though, Hegel zooms in on this apparent truism and highlights its fundamental inconsistency. Since being as such is nothing other than the absence of determinacy, to claim that everything *is* is really to claim *nothing at all*. The truth of being is really its own negation, that is, 'nothingness'. Yet the latter determination, when viewed more carefully, is itself purely negative. It can only be conceived as the opposite of something else, namely 'being'. This double criticism leads to the acknowledgement that 'being' and 'nothingness' are not, in fact, appropriate forms of conceptualizing reality. They are mere abstractions, artificially extracted from the continual flow of 'becoming'. This acknowledgement is the first in a series of sublations leading from the 'doctrine of being' to the 'doctrine of essence', and finally to the 'doctrine of the notion'. At the end of the Logic, the truth initially looked for in the notion of being finally emerges as the 'absolute Idea', a new and higher form of the subject-object relationship.

²⁵EL, 67 / HW 8, 92

²⁶EL, 60 / HW 8, 84

This new discovery brings the Logic to completion, but it also brings out its limited scope. In Hegel's words, the absolute Idea's ultimate fate is to abandon the subjective sphere and to release itself in 'the *externality of space and time*.'²⁷ The formal vocabulary of the Logic thus gives way to the concrete vocabulary of mechanics, physics and biology, and to the complex workings of the natural world. However, the Philosophy of Nature is not Spirit's final destination. Like the previous stages, it will be criticized and sublated, paving the way for the system's last and most sophisticated instalment. In the Philosophy of Spirit, instead of logical propositions or natural phenomena, the emphasis is placed at last on human beings, on their historical existence and on their ability to lead free and autonomous lives.

The previous description is only a brief summary, but it helps clarify the different ways in which Hegel's philosophical project can be deemed liberating. Returning to Will Dudley's formulation, it is now clearer why and to what extent the *Phenomenology* 'prepares the way for the free conceptual thinking that transpires in the logic, the philosophy of nature, and the philosophy of spirit.'²⁸ Just like Hegel's notion of *Wissenschaft*, his conception of liberation (*Befreiung*) can be interpreted in at least two fundamental ways: the first one refers to the *negative* process, described throughout the *Phenomenology*, whereby ordinary consciousness is freed from its natural cognitive imprisonment—the 'difference of consciousness' that haunts the different phenomenological stages; the second meaning of liberation refers to the *positive* process, narrated in Hegel's systematic works, whereby Spirit explores and actualizes its newly found freedom—from the subjective framework of the Logic all the way to the dialectical synthesis of 'absolute Spirit'.

Another way to make sense of this ambiguity is to return to Plato's dialogues, and particularly to the analogy drawn earlier between the *Phenomenology of Spirit* and the allegory of the cave. When discussing the possible outcomes of Plato's philosophical project, I distinguished between two forms of liberation. The first was purely *negative*: the cognitive purge promoted in the dialogues may lead, in the end, to an empty place. According to this hypothesis, the truth revealed by philosophy is to be nothing more than the *untruth* of human reason, and its ultimate reward a radical state of perplexity. The second possible outcome was *positive*, and more productive: instead of an empty place, the removal of the unwarranted presuppositions that dominate natural consciousness may reveal a new and truly rational outlook on reality. In this case, philosophy will bring forth not only the untruth of ordinary reason, but also a new and qualitatively different kind of truth. Recall, moreover, that this second alternative can also be construed in two different ways: this new outlook may be revealed instantly, as that which lies immediately beyond the doxastic realm, or it may be the beginning of a new philosophical quest, with its own challenges and breakthroughs. In Plato's allegorical language, the first of these scenarios corresponds to the one envisaged in the *Republic*, where the ascending prisoner ends up exiting the cave, seeing the sunlit world and knowing reality as it

²⁷SL, 753 / HW 5, 572

²⁸Dudley, *Hegel, Nietzsche and Philosophy*, 16

truly is. The second scenario requires a transformation of Plato's well-known image: instead of the journey's final destination, the exit from the cave is only the beginning of a new philosophical ascent, no longer *below* the earth but *above* it. The cave's subterranean path is followed by a new ascending road, along a hill or a mountain, towards a new cognitive altitude.

Considering the propaedeutic role ascribed by Hegel to the *Phenomenology* and his description of the system as the development of scientific truth, these two journeys seem to fit the two halves of Hegel's project. Whereas the phenomenological progression leads from the subterranean realm of self-deceit to the firm ground of 'absolute knowing', the system proposes a new elevation, in broad daylight, towards a new philosophical summit.

9.3 Hegel and Nietzsche

The *Phenomenology of Spirit*'s unique appeal, amply felt throughout the nineteenth and twentieth centuries, is due in large part to Hegel's tremendous faith in the power of human reason. Indeed, few other modern thinkers have placed such high hopes in philosophy's ability to reform human consciousness and free ordinary reason from its constitutive limitations. And this attitude is all the more remarkable since it can hardly be put down to a superficial engagement with human reason or to an indulgent diagnosis of consciousness' intrinsic limitations. If Hegel does stand out for his philosophical optimism, his work is no less notorious for its radical denouncement of human ignorance.

In the *Phenomenology*, these two aspects are mutually related. On the one hand, following a long lineage of critical thinkers, Hegel argues that the attainment of truth is conditional on the exposure of the intrinsic limitations of our ordinary mode of cognition. On the other hand, departing from the critical tradition, he claims that this negative procedure can have a positive outcome. He seeks to show that the untruth of ordinary knowledge is not simply untrue, but an undeveloped form of truth. Irrationality is itself bound to a rational principle and the critic's task is to identify this principle and follow where it leads.

Hegel's scepticism is thus a dynamic scepticism, where every particular criticism is the direct consequence of a previous criticism and the necessary antecedent of a new one. Human ideas and judgements are no longer taken up at random, as in Plato's dialogues, but placed within a global cognitive map, whose main itinerary allows for no detours. This ambitious strategy is grounded in two revolutionary convictions, which set the *Phenomenology* apart from all previous philosophical works. First, Hegel attempts to show that all aspects, dimensions and areas of human knowledge amount to different perspectives on the same cognitive whole. Whereas most of us tend to regard the fields of experimental science, morality, psychology, politics, religion or art as different worlds, whose aims and methods are irreducible to each other, Hegel regards them as different semblances of the same overarching Notion. The *Phenomenology*'s aim is to dissolve these distinctions and bring about a

truly global cognitive perspective, as rich and as complex as life itself. Second, Hegel also argues that this final perspective, due to the very form of the progression, is necessarily set for consciousness. If the different stages were merely added together, in the form of an aggregate, the progression might go on indefinitely, with no guarantee of success. But this is not the case. For Hegel, reason is inherently systematic, which means that the standpoint of science is not merely possible, but indeed logically necessary. Although consciousness may never leave the realm of untruth, the possibility of actual knowledge is no less real and within human reach.

With these two convictions, Hegel raised philosophy's old faith in reason and rationality to an unprecedented height. At the same time, however, his systematic solution can be seen as the final expression of this faith. Hegel was one of the last, if not the last great Western thinker to believe in the possibility of a truly sovereign mode of cognition, free from the chains of illusion and self-deceit. He was one of the last to hold that the contradictions that plague human life can be cured *through reason*, and not *against it*, or *beyond it*, or *in spite of it*. After Hegel, the philosophical canon shifted in an entirely new direction. His optimism was replaced by a fundamental mistrust of reason and science. From Schopenhauer's pessimism to Kierkegaard's existentialism, from Marx's materialism to Husserl's and Heidegger's 'pre-theoretical' turn, Hegel's insistence on harmony and reconciliation was inherited by a historical age at odds, perhaps more than any other, with the ideas of fragmentation, disharmony and conflict.

To borrow Karl Löwith's formula, the revolutionary break in nineteenth-century German philosophy coincided, to a great extent, with the rejection of Hegel and the emergence of different alternatives to systematic rationalism. In this context, Nietzsche is often pointed out as Hegel's most obvious countertype. This contrast is based on a number of well-known differences, which refer both to the content and to the form of their writings: whereas Hegel is ultimately an optimist, Nietzsche is famous for his 'Dionysian pessimism'²⁹; whereas Hegel is the quintessential systematic philosopher, Nietzsche is renowned for his dispersed aphorisms; whereas Hegel is a difficult and often abstruse writer, Nietzsche's prose is usually elegant and mordacious.

But Hegel's and Nietzsche's works also display important similarities. Firstly, both of their philosophies are built around the idea of science, or truth, and both can be interpreted as an attempt to answer the main challenge posed by the critical tradition. Nietzsche's idea of truth, just like Hegel's, is deeply influenced by the critical diagnosis set forth in Plato's dialogues. In his eyes, our usual cognitive perspective is flawed and self-deluded, burdened with unwarranted assumptions about ourselves and the world. These assumptions are for the most part automatic, embedded in human cognition and rarely called into question. Furthermore, these assumptions are not restricted to a specific brand of ideas or judgements, nor to a specific domain of experience. They contaminate the whole edifice of human reason and condition every aspect of human life.

²⁹GS, 236 / KSA 3, 622

Secondly, like Hegel, Nietzsche also regards reason's inner contradictions as a form of imprisonment and servitude. In line with the Platonic tradition, he also describes human life as a downgraded life and human beings as downgraded beings, reduced to a lesser version of their original selves, confined to a mode of cognition that compromises their power to choose and act freely. Moreover, he also construes philosophy's critical role as a liberating task, destined to reverse this state of affairs and re-establish humankind's original independence.

Thirdly, both Hegel and Nietzsche envisage this liberating movement as a cultural process, and both turn to the history of human reason to illustrate its main implications. What is more, the Hegelian and Nietzschean 'historiographies' of freedom share many important references. For one thing, both denote a clear fascination with the primordial innocence of the ancient Greek world, contrasted with the cultural degeneration of the modern Western world. In both accounts, the birth of Greek tragedy and the emergence of Socratic and Platonic dialectics constitute important conceptual innovations, with profound implications on the nature and scope of human freedom. Furthermore, Hegel and Nietzsche also highlight the cultural and conceptual role played by religion in the definition of freedom, and particularly by the rise and evolution of the Jewish and Christian traditions. Finally, both authors view Kant's critical project as a defining moment in the history of modern reason, with important repercussions on the definition of freedom.³⁰

However, these generic coincidences can only tell us so much. Hegel's and Nietzsche's historiographical appraisals are not only very different but, in many respects, antithetical. Hegel regards the phenomenological history of human cognition as the necessary path towards the actualization of freedom. On his view, each stage is necessary and liberating, provided it is correctly understood: although the harmony of Greek *Sittlichkeit* was irredeemably lost, its sacrifice was needed for the attainment of personal freedom; although primitive Christianity was a source of unhappiness, it played a decisive role in the development of modern reason; and although Kant's critical philosophy increased consciousness' self-awareness, its contradictions still call for a new *Aufhebung* and a more complete form of deliverance.

For Nietzsche, on the contrary, the idea that freedom is dependent on a systematic enquiry is seriously mistaken. Unlike Hegel, he rejects the basic premise of Plato's philosophical project and regards the pursuit of science not as a form of liberation, but as the primary cause of human bondage. Whilst Plato and Hegel anchored their criticism of human cognition in the notion of truth and set about denouncing the untrue nature of most human concepts and beliefs, Nietzsche argues that this grounding notion must itself be criticized. Its philosophical protagonism notwithstanding, the notion of truth is just a mere assumption, as groundless and unintelligible as all other human assumptions. Therefore, only a critical philosophy that is capable of acknowledging and overcoming this prejudice can actually be liberating. Whereas Plato and Hegel advocated a release *from untruth*, or a release

³⁰ See Sect. 1.3.

through science, Nietzsche reverses this general plan. He advocates a release *from truth*, or a release *from science*, and the embracement of a wider and freer standpoint.

Accordingly, Nietzsche's engagement with the history of human reason is very different from Hegel's. While Hegel viewed Greek dialectics, Christian morality and modern reason as necessary stages in the road to freedom, Nietzsche sees them as different and increasingly dogmatic incarnations of the scientific ideal, and hence as different and increasingly rigid forms of imprisonment. In his eyes, the long odyssey of human culture reflects, for the most part, the long evolution of the same grounding prejudice—a philosophical *parti pris* whose genealogy must be exposed and overcome.

Part III

Nietzsche's Metacognitive Project

Chapter 10

Nietzsche's Idea of Truth



Abstract In line with the imagery employed by Plato and Hegel, Nietzsche compares human beings to ‘involuntary slaves’ (KSA I, 74) or ‘fettered spirits’, ‘chained . . . to [their] pillar and corner’ (KSA II, 16). Descriptions of this kind abound in his writings, as do prophecies of a ‘great liberation’, or appeals to ‘the master’s privilege of the free spirit’ (KSA II, 16ff.). But although this kind of language evokes that of his predecessors, Nietzsche rejects Plato’s and Hegel’s liberating efforts. To be sure, he also believes humankind to be imprisoned or enslaved; moreover, he also locates the ultimate source of this imprisonment in our problematic relationship with truth. But whereas Plato and Hegel take truth as such for granted, and focus on how to attain it, Nietzsche questions the very need to think and choose in terms of truth and falsehood. In his eyes, the concern with truth is itself one of the dogmatic pre-suppositions that keep consciousness in bondage.

To clarify Nietzsche’s position, this chapter examines his unique conception of truth and his attempted reform of classical and modern epistemology. Focusing on published and unpublished writings, I analyse and discuss key Nietzschean motifs such as the ‘will to truth’ and the ‘will to illusion’, as well as the contrast between the *epistemic* ideal of the Platonic and Kantian traditions and the *sophistic* ideal Nietzsche seeks to revive and transform.

Keywords Nietzsche · Freedom · Truth · Illusion · Metaphor · Science · Wisdom

*Achtet mir, meine Brüder, auf jede Stunde,
wo euer Geist in Gleichnissen reden will:
da ist der Ursprung eurer Tugend.*¹

I have focused previously on the so-called allegory of the cave. Since Nietzsche is also fond of allegories, it seems appropriate to start this new string of chapters with one of his most famous images. The setting is no longer a subterranean cave, but a

¹TSZ, 57 / KSA 4, 99: ‘Pay attention, my brothers, to every hour where your spirit wants to speak in allegories: there is the origin of your virtue.’ (Translation modified.)

rocky mountain-top, and the lighting is no longer dim, but overflowing. It is the moment when Zarathustra, the prophet and 'free spirit', abandons his own cave, steps into daylight and addresses himself to the sun. After ten years of isolation, his heart has changed. He has grown weary of his wisdom and wants to pass it along. He has decided to return to the 'underworld', the world of human beings, and his future task is summed up in a formula that encapsulates Nietzsche's entire philosophical programme: to distribute his wisdom 'until the wise among human beings have once again enjoyed their folly, and the poor once again their health.'²

As so often in Nietzsche's writings, this description is filled with Platonic allusions. Apart from the cave and the sunlight, two of Plato's chief allegorical motifs, the whole picture is evocative of the *Republic's* allegorical situation. Zarathustra's descent into darkness, with the aim of freeing human beings from their natural ignorance, echoes the return of Plato's liberated prisoner to his original dwelling place, determined to share his discovery with his former companions. But just like the efforts of Plato's prisoner, Zarathustra's teachings are not taken seriously: his first speeches, directed at a crowd gathered in a marketplace, are greeted with incomprehension, laughter and hostility.³

In Plato's case, the allegory of the cave is meant to widen human beings' cognitive perspective and to unveil a possibility that is usually out of sight. Since all of us tend to regard the limits of our vision as the limits of the world, we are usually confined to our natural mode of cognition and bound to conceive difference as an immanent variation of the same cognitive alphabet. This peculiar form of imprisonment is the starting point of Plato's philosophical project. In the allegory of the cave and in other texts, he resorts to what I have called, following Hermann Fränkel, a Heraclitean thought pattern⁴: starting from the comparison between human life and that of the prisoners inside the cave, Socrates goes on to claim that these prisoners stand for *us*, that their life is *our* life. This declaration is the crucial moment of his narrative, the pivotal point at which description gives way to allegory. *We* are now the lower terms of a new comparison, whose higher term can only be envisaged negatively, as an utterly different kind of life.

Nietzsche's strategy is identical, and reiterated in different ways throughout *Zarathustra's* prologue. Apart from the aforementioned description, directly evocative of Plato's allegory, Nietzsche resorts to two other comparisons, both built according to the same allegorical scheme. First, he defines Zarathustra as 'an awakened one' (*ein Erwachter*) and other people as 'sleepers' (*Schlafenden*).⁵ This image is also Platonic⁶ and aims both to expose human beings' current awakeness as a *false* or *dream-like* awakeness and to suggest the possibility of a *genuine*

²TSZ, 3 / KSA 4, 11

³Cf. TSZ, 3ff. / KSA 4, 14ff. and *Re* 516e–517a. For Hans Blumenberg, Nietzsche's entire book amounts to 'a parody of Plato's cave.' (*Höhlenausgänge*, 617)

⁴See Sect. 1.2.

⁵TSZ, 4 / KSA 4, 12. Cf. BGE, 4 / KSA 5, 12, where Nietzsche defines the task of a new, post-Platonic breed of philosophers as 'wakefulness itself'.

⁶See Sects. 1.3 and 11.1.

awakeness, beyond the confines of ordinary consciousness. Second, in one of many references to Darwinism, Nietzsche compares human beings to apes, thereby introducing a new allegorical triad: just as apes are inferior to humans, so too are the latter inferior to a higher form of being, namely the 'overman' (*Übermensch*):

What is the ape to a human? A laughing stock or a painful embarrassment. And that is precisely what the human shall be to the overman: a laughing stock or a painful embarrassment.⁷

The term 'overman' is an allegorical term, whose meaning is not immediately intelligible. It has to be *derived* from the allegory's double comparison. Contrary to what is often suggested, Nietzsche's *Übermensch* is not a stronger or aggrandized version of our regular selves. As in Plato's allegory, what is here at stake is not a *quantitative*, but a *qualitative* upgrade: just as human cognition is qualitatively different from animal cognition, over-humanity lies beyond the sphere of human cognition and can only be attained by means of a radical reform of ordinary consciousness.

Nietzsche's Darwinian language is also meant to highlight the dynamic nature of this evolutionary scale. Ape, man and overman are not fixed or independent stages, separated by clearly defined boundaries. They are the moments of a living process, continually open to revision. Here, once again, Nietzsche's vision is strikingly Platonic: just as Plato defined human consciousness as the meeting point of two opposing forces (σῶμα and ψυχή, δόξα and ἐπιστήμη), Nietzsche compares mankind to 'a rope fastened between animal and overman'—and what is more, 'a rope over an abyss.'⁸ Following Plato, he highlights humanity's *hybrid* or *transitive* nature: human beings are permanently torn between 'under-humanity' and 'over-humanity'. They are not stable or self-sufficient beings, but incomplete and self-contradictory. Their life is 'a dangerous crossing, a dangerous on-the-way, a dangerous looking back, a dangerous shuddering and standing still.' It is 'a bridge and not a purpose', 'a *crossing over* and a *going under*' (*ein Übergang und ein Untergang*).⁹

In short, human beings are *no longer* animals and *not yet* overmen—they are something *in between*, *auf-dem-Wege*, μεταξύ. And this intermediate position is not stable or static. This point cannot be stressed enough: in Nietzsche's evolutionary scheme, human life is continually disrupted by a drive for self-transcendence; it is a permanent battle ground, as it were, where one can either stand or fall, rise to over-humanity or relapse into animality.

But these are still formal indications. Though it is now clearer *how* Nietzsche construes the relationship between *Mensch* and *Übermensch*, it has yet to be

⁷TSZ, 6 / KSA 4, 14. Apart from this passage's Darwinian overtones, Nietzsche is equally referencing Heraclitus's fragments DK 82 and 83, where human beings are also compared to apes. This is precisely one of the sources on which Fränkel based his notion of Heraclitean thought pattern.

⁸TSZ, 7 / KSA 4, 16

⁹TSZ, 7 / KSA 4, 16f.

determined *why* he draws such a distinction in the first place, *what Übermenschheit* consists in and *why* he regards it as a form of liberation. In the following chapters, I discuss the motivations, the causes and the consequences of this transformation. I start by considering its epistemological implications, linked to Nietzsche's redefinition of the notions of truth and untruth; afterwards, I deal with its moral implications, linked to Nietzsche's redefinition of the notions of good and evil, mastery and slavery, power and servitude; finally, I discuss Nietzsche's redefinition of freedom and his solution to Plato's original predicament.

10.1 The Will to Truth

In *Zarathustra's* prologue, the transition from humanity to over-humanity is said to depend on the attainment of *wisdom*—the *new* wisdom acquired during Zarathustra's retreat, contrasted with the *old* wisdom held by human beings. The dissemination of this newly found wisdom is supposed to make 'the wise among human beings enjoy once again their folly', and 'the poor enjoy once again their health.' At first glance, this initial declaration could also be interpreted in Platonic terms: 'to lead the wise to enjoy once again their folly' is a close enough definition of Plato's usual strategy, whereby human 'wisdom' is revealed as fundamentally illusory and human beings are led to acknowledge their natural 'folly'; in like manner, 'to lead the poor to enjoy once again their health' could be read, in Platonic vein, as the effort to 'cure' human beings' diseased standpoint through the acknowledgement of its own poverty. In either case, Zarathustra's revelation would coincide with Plato's main philosophical insight: human beings are 'of no account in respect to wisdom'¹⁰; their usual conviction regarding the truthfulness of their thoughts is dogmatic and unwarranted; in other words, their idea of truth falls short of *truth itself*.

There are, indeed, important parallels between Plato's and Nietzsche's characterizations of human cognition, which is worth considering in order to grasp the meaning and scope of the latter's 'Zarathustrian' revolution. To begin with, following Plato and a long philosophical tradition, Nietzsche highlighted the importance of the role played by the notion of truth in human life. In particular, he called attention to the fact that human thoughts and actions are never neutral or indifferent to their truth-value. On the contrary, they are supported by a global version of reality, usually assumed to be true. Just like the 'vital map' mentioned earlier, in connection with Plato's *Philebus*, or the 'cognitive alphabet' mentioned in connection with Hegel's *Phenomenology*, this version of reality offers a global diagnosis of one's existential situation: the diagnosis of what is at stake in one's life, the assessment of one's circumstances, the establishment of one's goals, etc. It provides the setting in which life usually unfolds and the insight needed for the success of one's decisions and actions.

¹⁰Ap 23a

Nietzsche refers to this phenomenon as the ‘will to truth’ and places it at the centre of his philosophical enquiries.¹¹ Human beings ‘will’ truth because they are dependent on a true diagnosis of what things are, of how they stand, of what they mean, etc. Yet he also points out that this ‘will’ is not absolute or uniform, but capable of different degrees of intensity. Even though all human beings are concerned with the truthfulness of their usual account of reality, this concern is not the only, nor indeed the strongest motivation behind their thoughts and actions. To ‘will truth’ is not necessarily to will it above all things, or at any price. While the ideal of complete or absolute truth might motivate philosophers and scientists—and even in their case only for a short while, if at all—it is certainly not what most people look for. In everyday life, the interest in truth is indeed pervasive, but not very demanding. Human beings rely on the notion of truth, but they are usually content with illusion; and although their adopted version of reality is usually assumed to be true, it is only as true as it needs to be in order to be believed. In most situations, most of us are either too busy or too lazy to get to the bottom of things. Unlike philosophers or scientists, we settle for a given definition or course of action without asking too many questions.

Most human ‘truths’ conceal the acceptance of a great deal of ignorance and self-deception. Nietzsche refers to this phenomenon as the ‘will to illusion’ and highlights its practical or existential utility.¹² If man were merely ‘a knowing animal’, ‘truth would drive him to despair and destruction.’¹³ In daily life, however, epistemic claims are usually subordinated to practical demands and the accuracy of one’s beliefs is conditional on how far they will help one pursue one’s practical interests. “‘Truths’ are demonstrated through their effects, not through logical proofs’, and the falsest judgements are sometimes the most indispensable to human life. ‘What is *true* and *what works* are regarded as identical.’¹⁴

In its customary form, then, human life is moved by two complementary drives. Usually, the ‘will to illusion’ has the upper hand: the practical ease felt in everyday life is a clear evidence that the problem of truth is never seriously raised. But even though one’s concern with truth is not very demanding, the ‘will to illusion’ is always grounded in a ‘will to truth’. If one’s ideas and judgements were not taken to be true, they would not be illusory, because an illusion is precisely an incorrect ascription of truth. Nietzsche’s contrast between ‘will to truth’ and ‘will to illusion’ does not refer, therefore, to a simple dualism, but to a cognitive gradation, ranging from a *pure* or *strong* concern with truth—a will to truth proper, free of illusion—to an *impure* or *weak* concern with truth—a ‘will to illusion’, with varying degrees of intensity.

¹¹ Nietzsche uses many different terms for this phenomenon: *Wille zur Wahrheit*, *Wille zum Wissen*, *Wahrheitstrieb*, *Wissenstrieb*, *Erkenntnistrieb*, *Pathos der Wahrheit*, *Hang zum Wahren*, etc.

¹² Nietzsche also uses many different terms for this phenomenon: *Wille zur Illusion*, *Wille zum Schein*, *Wille zur Täuschung*, *Trieb nach Glauben*, *Bedürfnis nach Glauben*, *Verlangen nach Illusion*, etc.

¹³ KSA 1, 760

¹⁴ KSA 7, 433. Emphasis added.

The 'will to illusion' is a specific mode of the 'will to truth'. Moreover, it is the mode that tends to dominate in everyday life. This general description is in many ways a recapitulation of Plato's doxastic model. But it is also the underlying model of every philosophical project grounded, in one way or another, in the idea of truth. We still find it in Kant, where the main issue at stake is the possibility of knowledge, and in Hegel, whose phenomenological project aims to reveal 'the actual knowledge of what truly is.'¹⁵ Generally speaking, all of these projects acknowledge the disagreement between the alleged truthfulness of human beliefs and the standard of truth they are supposed to fulfil. They are all based on different accounts of the cognitive scale ranging from truth to illusion.

Nietzsche argues that all previous philosophers, in their different efforts to expose the doxastic nature of human knowledge, were themselves subjected to the influence of a thoroughly doxastic presupposition—namely, the presupposition of truth's referential status. His critique is often articulated by means of visual metaphors, and particularly through the notion of *perspective*. Although the Platonic tradition and its subsequent developments aim to denounce ordinary consciousness' usual *lack of perspective*—that is, borrowing Schopenhauer's formula, its usual tendency to regard the limits of its vision as the limits of the world—they do so from an equally provincial or limited standpoint. By positing the ideal of truth ahead of their examinations, most philosophers ask us to 'think an eye which cannot be thought of at all, an eye turned in no direction at all, an eye where the active and interpretative powers are to be suppressed, absent, but through which seeing still becomes a seeing-something.'¹⁶ This eye is the 'eye of truth', to which nothing is hidden or invisible. And yet, Nietzsche argues, 'there is *only* a perspectival seeing, *only* a perspectival "knowing"; the *more* affects we are able to put into words about a thing, the *more* eyes, various eyes we are able to use for the same thing, the more complete will be our "concept" of the "thing", our "objectivity".'¹⁷

Habitually, the notions of perspective and perspectivism tend to be understood in one of two generic ways. To have a perspective on something might mean, firstly, that one views only an aspect or facet of a wider or more complex reality. One is not in possession of a global understanding of things, and this limitation may entail some form of self-deception. This, however, is not Nietzsche's meaning, for such an interpretation still assumes that there is a specific *truth* to be known, which is not yet fully grasped. Secondly, the notions of perspective and perspectivism might be taken to mean that there is no fixed or final version of things, and that a given object or event can only be known in partial and subjective ways. This interpretation, albeit closer to Nietzsche's meaning, is still not what he appears to have in mind. Instead of one truth, there are now several possible truths, but all of them are still grounded in the idea of cognitive adequacy. These different perspectives, despite their profusion, are still supposed to provide specific insights into what something *really is*, into its *true identity*, into its *actual meaning*, etc.

¹⁵ PS, 46 / HW 3, 68

¹⁶ GM, 87 / KSA 5, 365

¹⁷ Ibid. See, for ex., HAH, 9 / KSA 2, 20f.; GS, 213f. / KSA 3, 592f.; KSA 9, 466; KSA 11, 699ff.; KSA 12, 315; KSA 13, 373f.

In its most radical sense, Nietzsche's perspectivism proposes to overcome this general attitude. It is not directed at specific objects or realities whose true identity is more or less apparent, but at the *very notion of truth*. His point is that the tendency to think in terms of truth and untruth, or accuracy and inaccuracy, is itself a form of perspectival confinement. To adopt such an attitude is to reduce one's contact with reality to a cognitive perspective whose regional nature is not recognized as such. As Nietzsche puts it, to 'see' reality with the 'eyes' of truth and cognition is to forfeit the richness of one's original, unimpeded vision, and the various possibilities it allows.

To understand the full implications of this criticism, let us return once more to Plato's and Hegel's projects. For them, the problem of truth came down to the correspondence between subjective certainty and objective truth, *Begriff* and *Gegenstand*. However, Plato and Hegel did not simply confront human beliefs with an independent ideal of truth, brought in from the outside. They both argued that an epistemic standpoint can only be attained through an immanent critique of natural cognition. In Plato's case, this critique amounts to the *negative* procedure whereby Socrates attempts to 'purge' the minds of his interlocutors; in Hegel's case, to the *negative and positive* procedure whereby each phenomenological stage is sublated and replaced by a higher and more comprehensive one. But although Plato, Hegel and many others questioned the alleged truthfulness of human beliefs, they failed to question the value of truth itself. According to Nietzsche, they failed to recognize that truthfulness is not the only standard with which to assess the value of human beliefs.

This famous truthfulness that all philosophers so far have talked about with veneration: what questions this will to truth has already laid before us! . . . We paused for a long time before the question of the cause of this will—until we finally came to a complete standstill in front of an even more fundamental question. We asked about the *value* of this will. Granted, we will truth: *why not untruth instead?* And uncertainty? Even ignorance?¹⁸

The *idea of truth*—the 'in itself' which for Plato and Hegel represents the necessary backdrop of every act of cognition—is for Nietzsche the result of a prejudice. It is not necessary, but contingent; not the standard of measurement, but something more to be measured. In other words, the philosophical demand that every cognition be *true* cognition (whether in its stronger, epistemic variant, or in its weaker, doxastic one) does not follow from the definition of cognition, but from a specific and ultimately unjustified *attitude* towards cognition. It is the product of a peculiar kind of will, which tends to conceal its volitional origin and to pose as something essential and self-evident.

To venture beyond the realm of truth is the task of a new breed of philosophers, of which Zarathustra is the forerunner. Their mission is to replace the notion of truth with the more general notion of *value*. More precisely, their task consists in showing that false judgements may be *valid* judgements and true judgements *invalid* ones. Once the 'will to truth' is overcome, the truthfulness of a given judgement is no longer a valid argument for or against it. By then, human beings will have moved past this alternative and (presumably) learned to rely on new forms of valuation.

¹⁸BGE, 5 / KSA 5, 15

In this shift from *truth* to *value*, Nietzsche saw the key to a new and radically different outlook on reality. But what exactly does this transformation entail? And what exactly is its end result? A brief survey of Nietzsche's writings is sufficient to reveal the multifarious and often ambiguous nature of his answers to these questions. In line with his chosen methodology, he addresses this issue in a truly perspectivistic way, changing the tone, emphasis or angle of his description in order to make different points or to fit different contexts. Accordingly, his views on the 'transvaluation' of truth can be interpreted in at least three different ways, all of which are featured in his published and unpublished writings:

1. First, to overcome the 'will to truth' may amount to embracing a 'will to untruth', that is, a conscious or voluntary 'will to illusion'. While most of us are willing to accept various forms of illusion, we do not accept them *as illusions*, but precisely *as truths*. This does not mean, of course, that we do not deceive other people, nor that we are always against being deceived: dishonesty is an important part of human life, as is the aesthetic pleasure derived from fiction and fantasy. However, the kinds of deception we are ready to accept are usually very restricted. They refer to superficial elements of a much wider and more complex account of reality, whose truthfulness is usually taken for granted. When we go to the cinema, for instance, we may agree to being transported to a world of fiction, but this deception is very limited. Throughout the film, our idea of who we are, of where we are, of the identity of the people around us, etc., is not usually affected.

In order to 'consciously remain in untruth',¹⁹ we would have to relinquish our implicit faith in all of these judgements. The transformation envisaged by Nietzsche cannot be compared, therefore, to anything we know or have experienced. A life in which illusion—all illusion—were accepted and recognized *as illusion* would be very different from ours. It would resemble, perhaps, a fantastic dream, or a fantastic theatre play, where every incongruity is accepted as part of the plot. The falsehood of an idea or judgement would no longer be an argument against it, nor would truthfulness be synonymous with legitimacy. In this new world, contradiction would be as welcome as coherence, and encouraged whenever it proved advantageous.

But admitting that this transformation is possible, or even conceivable, its end result still presupposes the contrast between truth and untruth. To will illusion *as illusion* is still only to will *the opposite of truth*, i.e. untruth. Although truth has been denied its referential status, it is still employed, if only negatively, as a standard of valuation.

2. A second way of construing the argument amounts to recognizing the subsidiary nature of the contrast between truth and untruth vis-à-vis other, more essential forms of valuation. This theme is specifically evoked whenever Nietzsche speaks of the need to replace the 'philosophy of truth' with a 'philosophy of life', or 'health', or 'vitality', or 'power'. In *Beyond Good and Evil*, for example, the value of human judgements is said to depend not on their truth, but on 'how far [they] promote and

¹⁹HAH, 29 / KSA 2, 53f.

preserve life, how well [they] preserve, and perhaps even cultivate, the [human] type.’²⁰ In *The Gay Science*, Nietzsche extends this critique to philosophical judgements and calls for the replacement of conventional philosophers with philosophical *physicians*, capable of exposing philosophical ‘truths’ as ‘the unconscious disguise of physiological needs’²¹

As we will see below in more detail, this general argument is both descriptive and prescriptive. On the one hand, Nietzsche is claiming that human life is ultimately not about truth, but about health, growth, power, etc. The ‘will to truth’ is merely the disguise of a more primary existential drive—a ‘will to life’, or a ‘will to power’. On the other hand, Nietzsche is prescribing the acknowledgement and the correction of this mistake. He is advocating a return to life’s original or authentic form, beyond the artificial sway of cognition and truth.

Once again, the intended result of this transformation is the divorce between truth and value. Whenever false judgements prove empowering or life-promoting, they are accepted as valid judgements. Conversely, whenever an increase in accuracy entails a decrease in strength or vitality, the corresponding judgement or set of judgements is rejected. A life lived on these terms is indeed very different from our usual life, but the limitation found in the previous reading of Nietzsche’s argument is still present here, albeit in a subtler form. For even if this new solution is possible, or even conceivable, it still presupposes the notion of truth.

Admitting that one can become entirely indifferent to the truthfulness of one’s beliefs and care only about power or self-enhancement, this concern is still grounded in a global diagnosis of who one is, where one stands, what is at stake in one’s life, and so on. The pursuit of power and self-enhancement still presupposes a vital map in which power and self-enhancement occupy the central position. It is supported by the general conviction that ‘life is about power, or self-enhancement’—a conviction deemed *true*er than, say, the conviction that ‘life is about truth’, or about ‘pleasure’, or about ‘money’, or about anything else. In this general sense, the ‘will to power’ is still but a specific instantiation of the ‘will to truth’.

3. Finally, a third and more radical way of construing the argument is to take Nietzsche at his word and to envisage the overcoming of the ‘will to truth’ in *all its forms*. For human beings to break free from the idea of truth, they must overcome their usual need of a vital map. A will that is no longer regulated by the idea of truth can no longer ground its claims in an assessment of how things *are* and how they *are not*, of who one *is* and who one *is not*, of what is *at stake* in a given situation and what is *not at stake* in it. This extreme possibility raises a very difficult issue, concerning the criterion employed by this new form of consciousness. If the notions of truth and untruth cease to guide human judgements, which notions, if any, can take their place? Assuming that cognition requires some kind of normativity, what will the latter be based on? These and similar challenges call into question the very

²⁰BGE, 7 / KSA 5, 18

²¹GS, 5 / KSA 3, 348

possibility of Nietzsche's Zarathustrian revolution, as it is now being interpreted: is the notion of truth detachable from the notions of value, will, choice, power, and so on? And if so, to what extent? Is it possible to value something over something else without automatically deeming it *truer*?

All of these interrogations point to a new and uncharted territory. They seem to require an insight into a kind of experience that our usual standpoint is simply not equipped to conceive. Accordingly, Nietzsche often places the goal of his journey beyond the limits of humanity: perhaps the overcoming of truth, as he understands it, is indeed an 'overhuman' task, whose result cannot be appreciated from this side of existence.

10.2 Truth and Metaphor

The 'problem of truth' is arguably Nietzsche's central philosophical concern, revisited again and again throughout his writings. His solution to this problem, however, besides allowing for very different interpretations, is also characterized by a further and important ambiguity. In his *Zarathustra* and elsewhere, Nietzsche insists on the radical contrast between human beings' natural cognitive condition and a new, 'overhuman' condition. This contrast is meant to highlight the absolute incompatibility between our usual outlook on reality and the form it would have to assume if we were to leave behind our attachment to truth. There is no meeting point, no reconciliation, no common ground between these two existential attitudes: since human cognition is by definition incapable of grasping the kind of transformation Zarathustra prescribes, the only way to do so is to acquire an altogether new mode of being. To borrow Pyrrho's well-known formula, the only way to achieve this transformation would be 'to strip oneself completely of being human',²² and to put on a new existential garment.

In other texts, however, this evolutionary theme is replaced by a different approach to the problem of truth, namely a historical or genealogical approach. In these cases, Nietzsche's strategy is to put the 'will to truth' into perspective by arguing that human beings' usual attachment to truth is not a fixed cognitive trait, but an *acquired* one, whose nature and intensity have changed considerably over time. According to this line of argument, humanity's current cognitive condition is a degraded or decadent condition, deviated from its original form and in need of rehabilitation.

This genealogical strategy is first developed in a brief and incomplete text entitled *On Truth and Lying in a Non-Moral Sense*, written in 1873. The text's initial focus is once again the origin of truth's normative value: *When* did human beings' interest in truth first emerge? *Where* does it come from? *How* did it become such a generalized phenomenon? According to Nietzsche, the answers to these questions require a

²²Long and Sedley, *The Hellenistic Philosophers*, vol. 1, 14C

journey to the infancy of human history. Before the advent of communal life, when humans were still solitary creatures, truth and untruth were devoid of normative value. Primitive men and women were exclusively focused on their own survival and had no qualms in deceiving both other creatures and themselves. Their whole existence, Nietzsche argues, was based on a series of spontaneous, provisional and often self-contradictory beliefs, whose truth-value hinged solely on whether they were momentarily useful or useless, beneficial or baneful. Human intellect was fully subordinated to human life and reason counted solely as a life-enhancing tool, just like physical strength or dexterity.²³

At some point in time, due to ‘necessity and boredom’, this primeval solitude gave way to the creation of small communities, grounded in collective pacts of non-aggression.²⁴ Nietzsche’s description is thoroughly Hobbesian, and he even speaks of a *bellum omnium contra omnes*, replaced by a ‘peace treaty’ recognized and upheld by every individual. However, his primary concern is not the birth of political authority, nor the origin of social laws and institutions, but the historical evolution of the notions of truth and untruth.

In a pre-social environment, truth is a personal matter. Since there is no real communication, language is not yet needed and cognition is still made up of inarticulate beliefs, of spontaneous and changeable impressions. This is not to say, of course, that every human belief is equally valid, or that every human action is equally desirable. As Nietzsche will later write, ‘humans first placed value into things in order to preserve themselves.’²⁵ Before truth was elevated to a logical and cultural ideal, the value of human beliefs was entirely determined by the will to preserve and cultivate life. Therefore, although the primitive men and women imagined by Nietzsche are not yet social or scientific beings, their cognitive activity is already guided by a complex system of practical assumptions. They do not live in a state of permanent wonder, as though the world were created anew at each given moment. They are continually engaged in the process of *interpreting* reality and adapting their behaviour accordingly. Their intellectual activity consists in reducing the world to series of similarities and differences, in sorting what is important from what is irrelevant, what is advantageous from what is harmful.²⁶ They know, for instance, that a cloudy sky is the sign of an approaching storm and realize they must take shelter; they associate the shapes of different animals with different degrees of danger and different hunting techniques; they know that some seasons are warmer than others, that the air is usually colder at higher altitudes, that the sight of smoke is usually the sign of a fire, and so on.

²³Nietzsche’s criticism of the ‘will the truth’ can be interpreted in more radical and less radical ways. Considering my previous list, the present line of argument can be linked to the first and second interpretations, but not yet to the third: the primitive life-experience Nietzsche is documenting is already supported by a ‘will to truth’ in the more general sense, i.e. by a need for what I have termed a vital map. However, this primitive concern for truth is less intense or automatic than a modern and more robust ‘will to truth’.

²⁴See also HAH, 52f. / KSA 2, 95.

²⁵TSZ, 43 / KSA 4, 75

²⁶In other words, they are already driven by a ‘will to truth’ in the general sense.

In his eagerness to underline the contrast between primitive and modern reason, Nietzsche goes so far as to claim that for primitive men and women 'any conception of *natural* causality is altogether lacking.'²⁷ Although their understanding of reality is based on a series of practical assumptions, their ideas lack the force of actual laws. 'In regard to everything external . . . no conclusion can be drawn that something *will* be thus or thus, *must* happen thus or thus; it is [*they*] who are the more or less secure or calculable; man is the *rule*, nature is *irregularity*.'²⁸ In this primitive stage, all beliefs are inherently provisional and their interconnection is loose and variable. There is no explicit demand for logical coherence, no premium placed on intelligibility or verisimilitude. Although many things are expected to happen in a certain way rather than another, these assumptions are liable to change just as easily as they came to be adopted.²⁹

In a social environment, however, things are different. Personal assumptions must be externalized and shared with others. Their truth-value is no longer a personal matter, but a social one, and this change introduces the need for a coherent system of linguistic signs. Different things are accorded different names, and the latter must be recognized and adopted by an entire community. As a result, truth and untruth lose their former plasticity and become attached to specific linguistic combinations. They are henceforth the measure of the correspondence between a given thing or event and its appropriate designation.

This normative turn introduces for the first time the possibility of error, which can arise out of two different situations. A person can err involuntarily, by not knowing or forgetting how to use one or more linguistic signs, or voluntarily, by replacing a correct designation with an incorrect one. In the latter case, error gives way to deceit: the linguistic contrast between truth and falsehood leads to the more personal contrast between guilt and innocence, and the practical concern for truth will eventually pave the way to a *moral* concern for truth. But this new development

²⁷ HAH, 63 / KSA 2, 112

²⁸ HAH, 63 / KSA 2, 112f.

²⁹ It might be objected that this 'cognitive plasticity' is never as wide Nietzsche seems to think. Although some of the empirical connections that guide one's practical life are relatively weak, and thus more liable to be questioned—such as the assumption that a cloudy sky is the prelude of a storm, or that smoke is the sign of fire—there are other, much stronger connections whose validity is harder to challenge—such as the assumption that heavy objects have a tendency to fall to the ground, or that the sun rises anew every morning. Surely even a primitive human being is bound to realize that these things happen *every single time*. Nietzsche's point, however, is not that primitive human beings are not guided by empirical assumptions, but that they do not think of reality in terms of truth and untruth. Like modern men and women, they would be surprised if a heavy object, when dropped, would float in space instead of falling to the ground. But whereas for modern human beings this discovery would amount to the collapse of a whole system of beliefs, for primitive human beings it would simply prompt the endorsement a new and equally provisional constellation of empirical connections. In the latter case, the belief in 'gravity' would be revised simply because it would have lost its *instrumental* value, that is, because it would have proven useless or unreliable when applied to the fulfilment of a specific task. Nevertheless, this general explanation is still open to several objections, to which I will return in the following pages.

still lies ahead: when language is created, the notions of truth and untruth are still independent from both epistemological and moral concerns. Truth and honesty are not yet intrinsically valuable, nor desired for their own sake. Whereas human beings ‘desire the pleasant, life-preserving consequences of truth, they are indifferent to pure knowledge if it has no consequences’, and ‘actually hostile towards truths which may be harmful and destructive.’³⁰ Likewise, ignorance and dishonesty are not yet reproachable in themselves, but only insofar as they harm other people or the common good. ‘Human beings do not so much flee from being tricked as from being harmed by being tricked.’³¹ Hence, only when the failure to conform to the accepted linguistic standards proves dangerous or harmful is the ignorant scolded and the liar banned.

In this primitive scenario, no one wills truth *as such*. The notions of truth and untruth are contrived as a means of breaking the usual soliloquy of human experience and finding a common ground for sociality and communication.³² They are tools devised to assess the degree of correspondence between human experience and an arbitrary system of signs, established by decree and strengthened by habit. Consequently, to show that something is true or untrue is not to reveal what it is in itself, but simply to determine whether it conforms to the linguistic sign it was originally accorded.

However, Nietzsche’s argument is not simply about words, but about the actual things they stand for. The notion of truth and all its contradictions arise from the fact that words are never merely words, but the vehicles for specific theses about reality. Indeed, it is easy to see that the link between a given name and the thing it refers to is merely contingent: few people would argue that the English word ‘tree’, for instance, is in a better position to convey what a tree *really is* than, say, the German word *Baum*, or the Portuguese word *árvore*; these words amount to cultural and historical conventions, whose origin is ultimately just as arbitrary as that of any other word. The point, however, is that all of these words are usually uttered in the conviction that they refer to something real and definite, i.e. a tree καθ’αυτό, placed somewhere beyond the linguistic and conceptual realms. *Tree*, *Baum* and *árvore* are regarded as approximations to an actual empirical content, whose concrete truth is irreducible to this or that specific term.

Nietzsche’s aim is to show that this assumption is unfounded. And what is more, that it amounts to a complete subversion of the way human concepts are actually engendered. The notion ‘tree’ is born as a reaction to a nervous stimulation, which is then associated with an image, which is then translated by a specific sound. Nowhere in this process is a tree *in itself* to be found: the whole movement starts within the sphere of subjectivity and only afterwards does it spill over, so to speak, into the sphere of objectivity. As we grow accustomed to linking a specific stimulation with a

³⁰BT, 143 / KSA 1, 878

³¹Ibid.

³²Note the difference between Nietzsche’s and Hegel’s account, in which sociality is not a consequence, but the very basis of self-interest and self-identity.

specific image or sound, we tend to assume that the former is caused by something outside us, namely by an actual tree. What started out as a subjective impression ends up as a reified presence: like Pygmalion, we are surpassed by our own creations and led to believe their intrinsic truth.

As he retraces the genesis of the contrast between truth and untruth, Nietzsche arrives at one of the central themes of his early writings, namely the self-enclosedness of human knowledge and the impossibility of reaching beyond the sphere of subjectivity. Since all human concepts are self-referential, words do not grant access to reality *as such*—nor are they supposed to. What they do reveal is the relation of reality to the human mind. They document the different ways in which we are affected by this or by that nervous stimulus, the inner hierarchy of our perceptions, the properties of things we choose to highlight and the ones we fail to notice. In a way, human concepts are like mirrors: the image they display is our own image as seen through the different filters or lenses of everyday life.

To put it more plainly, the original source of the world's meaning is not the world itself, but our apprehension of it. *We* are the creators of the truths we put into words—the *artists* in whose image reality is forged. But this process is usually accompanied by a peculiar form of amnesia, favoured by time and habit, which makes us forget our creative role and regard truth and meaning as already given.

Nietzsche will later return to this same idea:

Because we have for millennia made moral, aesthetic, religious demands on the world, looked upon it with blind desire, passion or fear, and abandoned ourselves to the bad habits of illogical thinking, this world has gradually *become* so marvellously variegated, frightful, meaningful, soulful, it has acquired colour—but *we have been the colourists*: it is the human intellect that has made appearance appear and transported its erroneous basic conceptions into things.³³

To cognize is to create, and human reason is above all a creative process. But as names grow attached to the things they name, and as they are passed on from generation to generation, their initial arbitrariness falls out of view. Trees, animals or cars are no longer regarded as creative translations of subjective stimuli, but as *things in themselves* whose original or objective essence was translated into words and concepts. As a result of this change, human beings become convinced that their knowledge is wider and more efficient than it actually is. We come to believe that when we speak of trees, animals or cars 'we have knowledge of the things themselves, and yet we possess only *metaphors* of things which in no way correspond to the original entities.'³⁴

This use of the word 'metaphor' leads back to Hegel and his own concept of *Metapher*, or *Vorstellung*.³⁵ Much like Nietzsche, Hegel resorts to these terms in order to highlight the fundamental contrast between natural consciousness' supposed knowledge and its actual ignorance. As seen earlier, the Hegelian critique of natural

³³ HAH, 20 / KSA 2, 36f. Emphasis added.

³⁴ BT, 144 / KSA 1, 18. Emphasis added.

³⁵ Cf. EL, 30f. / HW 8, 44f.

cognition, developed in the *Phenomenology* and elsewhere, is set in motion by two complementary acknowledgements: first, the fact that human knowledge, in its or customary form, is usually taken to be clear and to provide a direct access to things in themselves; second, the fact that this assumption, when subjected to a closer examination, turns out to be groundless. In Hegelian terms, natural ideas are not clear or adequate, but mere *metaphors* of concepts whose actual content is out of reach. Yet Hegel's point is not simply that natural cognition falls short of actual knowledge, but that it is bound to *start out* as the appearance of knowledge. In other words, it is precisely because human ideas are intrinsically metaphorical, and act as the signposts of a meaning which is to be found elsewhere, that knowledge as such is possible. This centrifugal movement is not only the underlying basis of every natural idea, but also the starting point of the dialectical pathway towards knowledge.

Moreover, Nietzsche's and Hegel's descriptions are thoroughly Platonic in that they highlight both the fundamental gap and the fundamental bond that separates and unites human concepts and the meanings they are supposed to convey.³⁶ On the one hand, since each concept is part of an immanent cognitive alphabet, which extends to every domain of human experience, things in themselves are just as conceptual as any other thought-content. To think otherwise is to fall prey to a conceptual fiction and to mistake the sphere of subjectivity for the sphere of objectivity. On the other hand, this fiction is itself a necessary condition for the intelligibility of most, if not all human concepts. If a tree were merely a nervous stimulation, and not also an image, a sound and a physical presence, it would be utterly unintelligible, and hence of no practical use. For one cannot lean against a mere concept, nor climb it, nor stand on its branches, nor chop them up into wooden logs. It is only because the concept 'tree' is not usually experienced as a mere concept, but as something real and definite, that a tree is a meaningful part of our existential horizon. The verisimilitude of the concept 'tree', as that of any other concept we happen to make use of, is a basic requirement for the practical viability of our daily existence.

Furthermore, both Plato and Nietzsche argue that the practical need to assume the objective truthfulness of our concepts amounts to a peculiar form of amnesia. In other words, although human knowledge is entirely *metaphorical*, it is not lived or experienced as such. With time and practice, the metaphorical nature of our thoughts and ideas is gradually forgotten. Each of our concepts comes to be regarded as the *real thing*—or, at any rate, as an efficient or direct pathway into the real world.

³⁶In order to appreciate the various historical echoes here at stake, it is important to bear in mind that Nietzsche's use of the word *Begriff*—and my use of the word concept, following Nietzsche—is different from Hegel's. For the latter, *Begriffe* are usually opposed to *Vorstellungen* and refer to the *actual knowledge* which natural consciousness falls short of. Accordingly, Hegel often claims that ordinary knowledge is not yet *notional* or *conceptual* knowledge (*Begreifen*, *begreifendes Denken*), but the mere *appearance* thereof. For Nietzsche, on the other hand, concepts are the equivalent of Hegel's *Vorstellungen*, or Plato's δόξα. Hence, from Hegel's point of view, the parallel between his main argument and the Platonic and Nietzschean standpoints will only emerge if one replaces the words 'notion' or 'concept' with the words 'belief', 'representation', 'opinion', etc.

What, then, is truth? A mobile army of metaphors, metonymies, anthropomorphisms, in short a sum of human relations which have been subjected to poetical and rhetorical intensification, translation and decoration, and which, after they have been in use for a long time, strike people as firmly established, canonical, and binding; truths are illusions of which we have forgotten they are illusions, metaphors which have become worn by frequent use and have lost all sensuous vigour, coins which, having lost their stamp, are now regarded as metal and no longer as coins.³⁷

Time and habit have led human beings to forget about the metaphorical, or 'coin-like' nature of human concepts. As a result, men and women have come to regard trees, cars, their own bodies and all other experiential contents as things in themselves and to assume they have a clear enough knowledge of what those things *truly* are. Nietzsche's whole argument consists in exposing the contingent nature of this attitude and to counter its usual popularity. His aim is to unearth the possibility of a standpoint to whom things are so or so for merely *instrumental* or *aesthetic* reasons: a standpoint whose beliefs are no longer supported by conceptual or epistemic claims; a standpoint to whom it is ultimately irrelevant whether things are real in themselves or not, true or fictional, objective or subjective—in short, a standpoint whose metaphors are acknowledged and lived *as metaphors*.

Yet Nietzsche's overall argument, as it has been presented so far, is clearly ambiguous as regards the specific role played by words and concepts in the genesis of the contrast between truth and untruth. Firstly, as we have seen, he locates the origin of this contrast in a social compact, whereby 'a way of designating things is invented which has the same validity and force everywhere.'³⁸ This claim is then used to back the second part of the argument, concerning the origin of the human concern for truth. As individuals grow accustomed to ascribing specific words to specific things, they tend to forget that the former are mere conventions and come to believe in an essential link binding each word to the thing it refers to. But even if one accepts the idea that truth's normative value is linked to a social compact, it seems relatively obvious that the latter could never have arisen if some kind of normativity was not already in place. Inasmuch as one's pre-social life is already filled with assumptions and expectations, the pre-social world is already interpreted and hierarchized in countless ways, divided into specific shapes and meanings. The solitary men and women envisaged by Nietzsche may not need words, but they are just as capable of predicting specific events and ascribing (relatively) fixed meanings to things.

The question one should ask, then, is whether the will to truth is necessarily dependent on a social compact, or whether it can arise simply out of habit. In other words, whether pre-social individuals are not just as liable, with time and habit, to forget the subjective origin of their assumptions and mistake them for actual truths. This ambiguity might help explain why Nietzsche decides to complement his Hobbesian approach, in *On Truth and Lying*, with an alternative account of concept-formation, closer to classical empiricism. As the essay progresses, his

³⁷BT, 146 / KSA 1, 880f.

³⁸BT, 143 / KSA 1, 877

emphasis on *sociality* (and *words*) gradually gives way to an emphasis on *habit* (and *concepts*), and although the former line of enquiry is still necessary to explain the emergence of truth as a *moral* concept (an idea which is left for the most part undiscussed), the latter line of enquiry seems to take the upper hand, providing an account of the origin of the will to truth that is ultimately independent from the social compact mentioned at the outset.

But although this solution is more plausible, the role played by truth in the primitive scenario imagined by Nietzsche is still remarkably ambiguous. On the one hand, the whole argument hinges on the idea that truth and untruth are not natural, but acquired notions, and hence that human beings began, at some point in time, to think in terms of true and untrue things, ideas and states of affairs. On the other hand, however, if the will to truth is construed in general terms as the basic need to adopt or recognize a specific account of reality—a specific version of what or how things are, as opposed to what or how they are not—then it is hard to see how any human being could dispense with this kind of will. For surely primitive individuals, however ignorant, had to have some idea of the world around them in order to move about in it and envisage any goal or action whatsoever. In short, what Nietzsche's genealogical approach fails to make clear is what a life devoid of a specific account of reality, grounded in a basic set of cognitive references, would amount to, and whether it can be lived at all.

10.3 Science and Wisdom

According to Nietzsche, every truth is the result of a metaphor, and human cognition is the creative process whereby human life is translated into a 'mobile army of metaphors'. This definition is still somewhat vague, but two important aspects are already clear: first, the 'drive to form metaphors' is not regional or optional, but constitutive and transcendental. It 'cannot be left out of consideration for even a second without also leaving out human beings themselves.'³⁹ Secondly, it is as free and boundless as one's imagination. As the very word suggests, metaphors are much closer to artistic creations than they are to objective findings or scientific observations. Their origin is purely subjective and their ability to convey the meaning of a given reality or experience is ultimately an individual matter.

However, this last consideration only applies to *original* or *first-hand* metaphors, that is, to metaphors originally created by each individual to account for his or her own experiences. While Nietzsche's primitive men and women are indeed the creators of their own metaphors, and the sole judges of their value, modern men and women are usually not as creative or independent. They are born into a culturally charged world, grounded in several centuries of concept-formation. And though their experiences are in a way as personal as those of primitive men, they are no

³⁹BT, 151 / KSA 1, 887

longer mediated by original metaphors. Their world is already labelled and documented, explained and hierarchized. Since cognition is now chiefly a matter of correspondence, objects and events are no longer as open to different meanings or interpretations: there is a *true* concept (or set of concepts) for every empirical content, and every other combination is deemed *less true*, or altogether *untrue*. Every failure to conform to this system is usually put down to ignorance, irrationality or artistic license: if one applies the term 'chair' to a table, for instance, or claims that human beings are 'four-legged animals', one's words will have a poetic meaning, or none at all.

As shown above, Nietzsche regards this rigidity as the result of a collective amnesia, which led to the elimination of life's original malleability. In his own words,

Only by forgetting this primitive world of metaphor, only by virtue of the fact that a mass of images, which originally flowed in a hot, liquid stream from the primal power of the human imagination, has become hard and rigid, only because of the invincible faith that *this* sun, *this* window, this table is a truth in itself—in short, only because man forgets himself as a subject, and indeed as *an artistically creative* subject, does he live with some degree of peace, security and consistency; if he could escape for just a moment from the prison walls of this faith, it would mean the end of his 'consciousness of self' [*Selbstbewußtsein*].⁴⁰

The idea that life's practical viability is conditional on a stable and relatively dogmatic system of truths is fairly intuitive. As Nietzsche himself concedes, if human cognition were to relinquish its normative dimension, and if all human meanings were suddenly open to revision, there would be no peace, security or consistency. If human imagination were to regain its fluidity, the notions of law and causality would be called into question and experience itself would be radically transformed.

But one does not have to imagine such an extreme scenario in order to appreciate the practical utility of conceptual dogmatism. In fact, if only half of human concepts were to lose their normative value—nay, if only one hundredth, or even one thousandth of all human concepts were to suffer a similar fate, our usual outlook on reality would already be significantly altered. If we ceased to rely on the implicit truthfulness of tables, windows, trees or cars, the specific role they play in our lives would be transformed. The practical ease with which we sit at a table, look out of a window, climb a tree or jump into a car stems from our ingrained belief in the objective truth of these concepts. If tables, windows, trees and cars would appear to us as subjective creations, and no longer as things in themselves, the vast array of qualities, uses and expectations we have come to associate with them would also be transformed. They would no longer blend into our existential landscape as they normally do, but stick out as new, unfamiliar and potentially unrecognizable elements.

⁴⁰BT, 148 / KSA 1, 883f. As noted in Ronald Speirs' translation, the word *Selbstbewußtsein* can mean both 'self-consciousness' and 'self-confidence', or 'self-conceit'.

Furthermore, if the objective truth of these and other concepts were radically questioned, the resulting transformation would not be limited to specific elements or regions of our cognitive horizon. It would affect, in one way or another, our entire understanding of reality. For unlike what we are usually led to assume, concepts are never atomic or elementary determinations: they cannot be reduced to simple cognitive events, nor considered in isolation. Their meaning can only be grasped in context, as part of a complex semantic web whose ramifications extend to the whole conceptual sphere.⁴¹ Returning to my previous example, a tree is always much more than a *simple* or *actual* tree—and this is so for at least three main reasons: first, because the concept ‘tree’, in order to be understood, requires the simultaneous consideration of an endless series of related concepts, without which the concept ‘tree’ would not be intelligible (namely ‘solidity’, ‘unity’, ‘identity’, ‘plant’, ‘life’, etc.); second, because a tree is never simply an individual or isolated tree, but a universal genus, of which each an individual tree is but a concrete instantiation; third, because the concept ‘tree’ is itself necessarily implicated in the meaning of many other concepts (namely ‘trunk’, ‘forest’, ‘orchard’, ‘willow’, etc.).

This and similar examples demonstrate what Nietzsche argues by dint of architectural images: namely, that human concepts are never loose or independent elements, but the constituents of a vast and intricate cognitive edifice, where every brick or beam has its proper place, and where all contribute in a specific way to the harmony of the ensemble.⁴² By ‘dissolving images into concepts’, or by ‘sublimating sensuous metaphors into a schema’,⁴³ mankind managed to escape the fleetingness of sense perception and the solitude of subjectivity. The normative turn from arbitrariness to abstraction enabled ‘the construction of a pyramidal order based on castes and degrees, the creation of a new world of laws, privileges, subordinations, definitions of borders, which now confronts the other, sensuously perceived world as something firmer, more familiar, more human, and hence as something regulatory and imperative.’⁴⁴

This vast and admirable construction has been the source of much more than ‘peace, security and consistency’. It was and still is the source of marvellous feats, the guiding force behind all sorts of practical and theoretical exploits. Throughout millennia, this conceptual dogmatism has helped humans counter their natural limitations, heighten their physical and intellectual powers, extend and improve their lives. In light of all this, Nietzsche’s claim that ‘conceptual faith’ amounts to a form of imprisonment is not as intuitive. For even if something has been lost, if imagination has indeed been stifled by the dogmatism of conceptual reason, a lot

⁴¹ A locus classicus for this idea is *So* 251ff., where Plato argues that human concepts are, by definition, *relational* determinations, naturally bound to *participate* in (κοινωνεῖν, μετέχειν), or *comingle* with (μυγῆναι), other determinations.

⁴² For a discussion of Nietzsche’s architectural metaphors, see Kofman, *Nietzsche et la métaphore*, 89–117.

⁴³ BT, 146 / KSA 1, 881

⁴⁴ BT, 146 / KSA 1, 881f.

more seems to have been gained. And surely the primordial fluidity of human creativity, albeit free from the burdens of consequence and coherence, was itself a much more damning form of imprisonment. In the absence of stable definitions, of fixed habits and laws, humans were doomed to be 'swept away' by sudden impressions. Like explorers travelling through an ever-changing landscape, they were unable to rest or settle down. Their freedom—if one can call it so—was too wide to be liberating.

What should we make, then, of Nietzsche's praise of this primitive freedom? He does not seem to be advocating a simple return to ignorance, or a complete departure from the sphere of conceptual reason. Such a project would not only be self-contradictory—its very formulation is intrinsically conceptual—but also, in all probability, utterly unfeasible. As we have seen earlier, with the aid of Plato and the *Philebus*, human cognition is irreducible to the spheres of sense perception (αἴσθησις) and memory (μνήμη).⁴⁵ Every act of cognition, however simple, already entails a specific diagnosis of reality—a silent claim (or constellation of claims) about how things are, what they mean, and how they relate to one another. And this diagnosis is by definition a *normative* diagnosis: it presupposes the distinction between adequate and inadequate representations of objects and events. All of this is of course irreducible to the mere storing and combining of sensations. It is the work of a thoroughly empirical standpoint, whose cognizing is already a form of distinguishing, sorting out, comparing and judging.

In addition, we have also seen that an empirical standpoint such as ours is never stuck in the present moment. Our diagnosis of reality is always a proleptic diagnosis, whose aim is just as much to determine what things are (and were) as it is to anticipate what they will be. To cognize is thus not simply to describe or judge, but to plan and anticipate; to harbour expectations, based on one's current and past experiences, about how things are supposed to evolve. And this tendency is the direct cause of another of our standpoint's defining features: namely, the propensity, heightened by time and habit, to transform our empirical observations into fixed laws, endowed with objective truth. In Platonic terms, this propensity speaks to the *doxastic* nature of human cognition; in Nietzschean terms, it reflects the historical tendency to forget human cognition's metaphorical foundation.

If human beings' native standpoint is not sensuous or mnemonic, but empirical (or metaphorical), Nietzsche's criticism of conceptual reason must also be understood and interpreted within the bounds of experience. More specifically, if his liberated standpoint is to be at all conceivable, it must be located somewhere within the narrow margin that separates ἐμπειρία from δοξάζειν—that is, it must be able to take its own metaphors *seriously enough* to ensure the world's practical viability, and yet *not as seriously* as ordinary consciousness usually does. This is the compromise Nietzsche seems to have in mind when contrasting the so-called 'rational human being' (*der vernünftige Mensch*) and the 'intuitive human being' (*der*

⁴⁵ See Sect. 2.1.

intuitive Mensch).⁴⁶ Whereas rational human beings answer to the scientific imperatives of coherence and objectivity, and tend to forget themselves as creative subjects, intuitive human beings are fully aware of their role as creators of truth and meaning. They also rely on empirical concepts and observations, but their attachment to objectivity is of a different nature. Their idea of value is not confined to the alternative between truth and error, or to the contrast between a world of appearance and a world in itself. Concepts are for them mere tools, which can be used in certain occasions and discarded in others. They bring out certain aspects of reality, just as specific colours or sounds are used by artists to produce certain visual or acoustic effects. Inasmuch as the value of these aspects of reality is irreducible to a logical or objective criterion, they are the opposite of science's cool abstractions: their meaning is as varied and incoherent as life itself, and although their infinite palette also includes scientific reason, the latter is but one amid a multitude of different cognitive hues.

This freedom is unknown to rational human beings, for whom metaphors are no longer lived or recognized as such. Nevertheless, the aesthetic foundations of human experience are still available to them in the form of art and artistic creation, where 'the intellect, that master of pretence, is free and absolved of its usual slavery.'⁴⁷ Through art, scientific individuals are offered a glimpse of their original identity and led to abandon, if only momentarily, their usual dogmatism.

The intellect has now [through artistic creation] cast off the mark of servitude; whereas it normally labours, with dull-spirited industry, to show to some poor individual who lusts after life the road and the tools he needs, and rides out in search of spoils and booty for its master, here the intellect has become the master itself and is permitted to wipe the expression of neediness from its face. Whatever the intellect now does, all of it, compared with what it did before, bears the mark of pretence, just as what it did before bore the mark of distortion.⁴⁸

The key to understanding Nietzsche's point lies precisely in the difference between the voluntary *pretence* of artistic creation and the involuntary *distortion* usually involved in empirical conceptualization. While both cognitive attitudes are grounded in a thoroughly subjective procedure, only the first one is able to acknowledge and take advantage of the true power of human imagination.

That vast assembly of beams and boards to which needy man clings, thereby saving himself on his journey through life, is used by the liberated intellect as a mere climbing frame and plaything on which to perform its most reckless tricks; and when it smashes this framework, jumbles it up and ironically re-assembles it, pairing the most unlike things and dividing those things which are closest to one another, it reveals the fact that it does not require those makeshift aids of neediness, and that it is now guided, not by concepts but by intuitions.⁴⁹

⁴⁶See BT, 150ff. / KSA 1, 886ff. Translation modified.

⁴⁷BT, 151 / KSA 1, 888

⁴⁸BT, 152 / KSA 1, 888

⁴⁹Ibid. The 'boards and beams' needed for one's journey through life evoke the metaphorical raft imagined by Simmias in *Phd* 85d, also needed to 'sail through life'.

But what does a life guided by intuitions look like? How does it differ, in practical terms, from the lives of modern, 'needy' individuals? For Nietzsche, the answer can be found in human history. 'Intuitive human beings'—viz. 'free individuals', 'artistic individuals', 'liberated intellects', etc.—are not just a conceptual or philosophical anti-type. They are the product of a lost cultural era, whose cognitive freedom is no longer intelligible or within reach. This lost world is that of Zarathustra, where reason and imagination, earnestness and playfulness, lucidity and madness went hand in hand. But it is also the world of ancient Greek mythology, where art and imagination used to rule over objective truth, and where each word and gesture were accompanied 'by pretence, by the denial of neediness, by the radiance of metaphorical visions, and indeed generally by the immediacy of deception.'⁵⁰ To quote another of Nietzsche's formulas, the 'Greeks were superficial—*out of profundity!*'⁵¹ They drifted from logic to myth with the ease of artists, and this ingenuity was also visible in their unique cosmology: unlike the world of modern individuals, entirely made up of 'things in themselves', the Greek world was peopled with gods and other fantastic beings, subject to magical influences and hidden forces. It was not a mere depository of objective truths, but the living mirror of human passions, fears and aspirations.

The contrast between the ancient Greek world and the modern world is a recurring Nietzschean motif, explored in different ways throughout many different works. It plays a particularly important role in the early writings on Greek tragedy, where the origin of classical theatre is famously located in the historical synthesis of the so-called Dionysian and Apollonian principles.⁵² In his 1886 preface to *The Birth of Tragedy*, Nietzsche points out that his analysis of Greek tragedy is already an attempt to solve 'the *problem of science* itself'. In it, he claims, science is 'grasped for the first time as something problematic and questionable.'⁵³ And the main novelty of his approach lies in his adoption of an artistic rather than a strictly epistemological approach to the problem of science. He has realized that 'the problem of science cannot be recognized within the territory of science.'⁵⁴

⁵⁰ BT, 152f. / KSA 1, 889. In this and similar passages, Nietzsche claims a) that ancient Greek men and women were naturally artistic beings, who favoured the *aesthetic* over the *concrete* aspects of life; and b) that the kind of ideas and values cherished by ancient Greek society can no longer be understood or shared by modern human beings. Both of these ideas are familiar motifs of German Romanticism's revival of ancient Greek culture, poignantly conveyed by Schiller's 1788 poem *The Gods of Greece* (*The Poems of Schiller*, 65): 'Ah! how diff'rent was it in that day! . . . / When the magic veil of Poesy / Still round truth entwined its loving chain— / Through creation poured life's fullness free, / Things then *felt*, which ne'er can feel again.'

⁵¹ GS, 9 / KSA 3, 352

⁵² This group of works includes *Zwei öffentliche Vorträge über die griechische Tragödie* (1870), *Die dionysische Weltanschauung* (1870), *Die Geburt des tragischen Gedankens* (1870), *Sokrates und die griechische Tragödie* (1871) and the well-known *Die Geburt der Tragödie* (1872, republished in 1886), which incorporates material from the previously mentioned texts. For an overview of Nietzsche's engagement with Greek philosophy and culture, see Müller, *Die Griechen im Denken Nietzsches*.

⁵³ BT, 4f. / KSA 1, 13

⁵⁴ BT, 5 / KSA 1, 13

Nietzsche proposes to carry out two main tasks, namely ‘looking at science through the prism of the artist’ and ‘looking at art through the prism of life.’⁵⁵ The first of these tasks is also attempted, as we have just seen, in his 1873 opus-cule *On Truth and Lying in a Non-Moral Sense*. Indeed, both texts display important similarities: in the *Birth of Tragedy*, Nietzsche also draws a clear distinction between the aesthetic freedom of ancient Greek culture and the scientific objectivism of modern culture; he also opposes a primeval ‘kingdom of wisdom’, where fables and metaphors were still cherished and taken seriously, to the ‘logical despotism’ of conceptual reason; and, moreover, he also contrasts the freedom and the sovereignty of the Greek standpoint with the slavery and the servitude of the scientific attitude.

The work’s second task, however, adds a new element to Nietzsche’s narrative, which is only briefly addressed in the second section of the 1873 opus-cule, left unfinished. In addition to the *epistemological* issue of ‘where the will to truth comes from’, discussed throughout the first section, Nietzsche also focuses on the *existential* issue of ‘whether life is worth living’, and on the role played by truth and untruth in this debate. The importance of this new theme is also reiterated in the 1886 preface, where the whole text is said to be about the ‘value of existence’, and in the new subtitle which accompanies the work’s second edition: ‘Hellenism and Pessimism’.

But how does the contrast between optimism and pessimism—or, as Nietzsche also puts it, between the *affirmation* and the *negation* of life—relates to our previous considerations regarding truth, freedom and metaphor? We have seen that the starting point of human cognition is, according to Nietzsche, a transcendental ‘drive to form metaphors’. Each human being, naturally compelled to give shape to his or her subjectivity, is led to translate his or her inner life into a palpable existential setting. The human intellect ‘makes appearance appear’ and endows it with a specific form, a specific texture, a specific structure. Furthermore, we have also seen that this original drive can result in different degrees of freedom, depending on the extent to which one is aware of oneself as an artistically creative subject. Whereas ‘intuitive human beings’ avail themselves of all kinds of cognitive resources, and are thus able to express their subjectivity more freely and more completely, ‘rational human beings’ are locked within the ‘prison walls’ of objectivity.

Near the end of the *Truth and Lying* essay, Nietzsche goes on to argue that these two cognitive approaches are grounded in two different existential attitudes. Both kinds of human beings desire to ‘rule over life’, but they attempt to do so in different ways. Rational individuals are diligent, conscientious and thoroughly obedient creatures, whose only concern, like the ant in Aesop’s fable, is to provide for the future and guarantee a safe and regular life. They put their faith in empirical concepts and abstractions in order to avoid pain and misfortune, and their efforts usually pay

⁵⁵ BT, 5 / KSA 1, 14

off. However, their obedience is unable to bring them real happiness. It is only a means of defence, a way of minimizing the uncertainty of an otherwise unbearable life. By contrast, intuitive individuals are carefree, spontaneous and exuberant. They are not intellectual, but aesthetic beings, and their whole life is lived in pursuit of joyous visions, feelings and effects. They 'only [acknowledge] life as real when it is disguised as beauty and appearance', and reap directly from their intuitions 'not just protection from harm but also a constant stream of brightness, a lightening of the spirit, redemption and release.'⁵⁶ However, they suffer more and more severely than rational human beings, as they are less able or less willing to learn from their past misfortunes. But their suffering is as just as unreasonable as their pleasure, and just as free from the constraints of objective truth.

The *Birth of Tragedy* is built around a similar motif, but Nietzsche's main argument is more clearly grounded in a pessimistic diagnosis of human life. Inspired by Schopenhauer, he argues from the beginning that life is *not* worth living, but that it can be 'redeemed' or 'justified' as an *aesthetic* phenomenon and that this discovery is the greatest achievement of ancient Greek culture. Because of their uniquely artistic nature, the Greeks were also uniquely sensitive creatures, and hence uniquely aware of the horrors of human existence. Their entire worldview, Nietzsche argues, can be interpreted as a reaction to the so-called 'wisdom of Silenus', that is, the acknowledgement that being born is not a blessing, but a burden, and that being alive is ultimately worse than *not to be* at all.⁵⁷ To cope with this truth, the Greeks took refuge in a realm of disguise and appearance: they replaced the world's primitive hostility with a realm of beauty and divinity, filled with sacred beings and sublime shapes. With time, this *Apollonian* instinct for beauty succeeded in casting its veil over every aspect of human existence, rendering life pleasant and worthwhile. But this love of disguise is utterly irreducible to the idea of deception or distortion, as it is understood by modern reason. This veil of beauty is not to be regarded as a curtain drawn over a world in itself, whose objects were somehow *truer* or *more real* than the objects of Greek imagination. For this kind of reading is already contaminated by modern reason's objectivist framework. As Nietzsche continually points out, Greek gods and myths are not false because they are not true, in the objective or scientific sense of the word. They are metaphors *lived as such*, and indeed cherished *because of it*. Unlike modern individuals, the ancient Greeks revelled in the act of pretending. They did not divide the world into real and unreal objects, but into tasteful and distasteful images, charming and dull metaphors. Their chief concern was the extent to which each new image was able to impress and delight, to infuse pleasure and joy.

In like manner, Nietzsche's central contrast between the Apollonian and Dionysian principles must not be construed as an intellectual or scientific contrast. Neither of these stances claims to be truer than the other one, nor to offer an image of the world as it actually is. What is at stake is once again an aesthetic opposition, which

⁵⁶BT, 152 f. / KSA 1, 889

⁵⁷BT 22ff. / KSA 1, 34ff.

must be understood from within the metaphorical sphere. The Apollonian element is the element of creativity and pretence, responsible for ‘making appearance appear’, while the Dionysian principle stands for the destruction of appearance and the elimination of the metaphorical plane. Dionysos is the god of intoxication and madness: under his spell, the principle of individuation is threatened and subjectivity itself is jeopardized.

It must be noted, however, that the *Birth of Tragedy*’s initial argument, as summed up in the previous paragraphs, is not Nietzsche’s final word on the relationship between art, truth and life. In his 1886 preface, he distances himself from Schopenhauer and his pessimistic worldview. During the more than 10 years that separate the initial text from this self-critical preface, Nietzsche altered his views on optimism and pessimism and came to regard the latter as a specific symptom of modern reason. In most of his later writings, the grandeur of ancient Greek culture is located precisely in the readiness to accept and embrace life, whereas pessimism and nihilism are construed as the preeminent signs of modern European decadence. This change of strategy requires a reformulation of the *Birth of Tragedy*’s initial diagnosis and the abandonment of the original emphasis on the ‘wisdom of Silenus’. In an attempt to harmonise his past and present views, Nietzsche opens the preface with a somewhat factitious distinction between different kinds of pessimism: a ‘pessimism of *strength*’, indicative of physical health and vitality, and a *weaker* sort of pessimism, associated with Schopenhauer and modern culture as a whole.⁵⁸

The importance of this inconsistency will become more apparent in the next chapters, in the sections dedicated to the relationship between life, health and power. My present purpose is to highlight the continuity between Nietzsche’s views on truth and metaphor, as expounded in the *Truth and Lying* essay, and the contrast between the Dionysian and Apollonian principles. In both cases, the decline of Greek creativity is attributed to a peculiar kind of amnesia, whereby the freedom of imagination is replaced by the prison of conceptual reason. But whereas in the first case this movement is put down to time and habit, in the second case it is documented in a different and more detailed manner.

It all comes down to the peculiar balance between Apollo and Dionysos. Although these two principles represent opposite forces, they are intrinsically dependent on one another. Just as the violent excesses of the Dionysian *Trieb* must be tempered by the serenity of the Apollonian *Trieb*, so too is the latter only truly meaningful as a reaction to the former. Dionysos acts a counter-weight to the world of metaphor, reminding the subject-artist of his or her own condition. In Nietzsche’s own words, Dionysos reveals ‘the playful construction and demolition of the world of individuality as an outpouring of primal pleasure and delight.’⁵⁹ This central idea is further illustrated by the evocation of the Heraclitean fragment where

⁵⁸Nietzsche returns to this distinction in *The Gay Science*, where he contrasts the *romantic pessimism* of Schopenhauer and Wagner, born from an ‘impoverishment of life’, and the *Dionysian pessimism* of ancient Greece, born from a ‘superabundance of life’ (GS, 234ff. / KSA 3, 619ff.).

⁵⁹BT, 18 / KSA 1, 30

the power that shapes the world is compared to the action of a child playing on the beach, building up piles of sand only to knock them down again.⁶⁰ What is important in these images is the fact that *both* moments, creation and destruction, are equally essential and intrinsically related. If the creative moment were ignored, subjectivity would be eliminated and Dionysos would reign supreme; if the destructive moment were neglected, the metaphorical world would grow rigid and out of proportion. And while the first of these alternatives might correspond to some embryonal stage of human reason, wherein subjectivity was not yet fully formed, the second one is precisely what took place, according to Nietzsche, in the period that led to the demise of ancient Greek culture.

The course of the argument is well known: as Greek art became more intellectualized, the Apollonian element grew increasingly prominent. In the case of tragedy, it was finally chased by Euripides from the theatre stage. But what happened to tragedy was merely a symptom of a much wider transformation, which struck the very heart of Greek culture. Free from the Dionysian menace, human metaphors grew higher and sturdier. Piles of sand were replaced by wooden towers, by large stone bulwarks—and soon the world of metaphor was no longer recognizable. Apollonian visions crystallized into cold abstractions and aesthetic freedom gave way to epistemic truth. The artistic drive to impress and delight was replaced by the scientific wish to pierce the veil of illusion and reveal the world *as it really is*.

Nietzsche's historical description is now more precise. The death of metaphor is linked to a specific cultural event, brought about at a specific juncture of human history. Although the demise of tragedy, poetry and Greek art in general was a gradual process, it was not due simply to time and habit. In this new account, the birth of scientific reason coincides with the rise of one man—Socrates—and the endorsement of a new form of culture—Platonic culture.

⁶⁰Ibid. Fragment 52 in the Diels-Kranz edition. See also GM, 58 / KSA 5, 323f.

Chapter 11

Philosophy in Chains



Abstract This chapter retraces Nietzsche's genealogical account of the 'will to truth'. Like Hegel, in whose writings the pursuit of truth and science takes the form a long conceptual progression, with historical and cultural implications, Nietzsche also regards the conquest of truth as a long progression, but leading in the opposite direction: the history of Western thought is presented as the history of a cognitive decline, from the aesthetic freedom of the pre-Socratic world to modern culture's increasingly unhealthy fixation on truth, logic and science.

While Sect. 11.1 discusses Nietzsche's characterization of Socrates as the destroyer of classical tragedy and the founder of modern scientific culture, Sect. 11.2 analyses his views on the moral and religious implications of Socrates' and Plato's scientific revolution. In his critique of morality, Nietzsche explicitly appropriates the master-slave imagery employed by Plato and Hegel. This issue is taken up in Sect. 11.3, which focuses on the notions of 'master morality' and 'slave morality' and highlights their practical and epistemological implications. Finally, Sect. 11.4 offers a discussion of the final stage of Nietzsche's genealogy, i.e. the growing nihilism of modern reason, grounded in the 'ascetic ideal' that pervades the history of Western philosophy.

Keywords Nietzsche · Plato · Socrates · Judaism · Christianity · Freedom · Slavery

And then Socrates came along—and his influence 'spread out across all posterity to this very day, and indeed into the whole future, like a shadow growing longer in the evening sun.'¹ Anyone who has but a superficial acquaintance with Nietzsche's works will know how important a role Socrates plays in them. Whether merged with Plato or standing on his own, he is the object of a variety of different and often contradictory reproaches: he is called a demon and a slave, a genius and a charlatan, a naïve optimist and a conceited despot. But in neither of these roles is the importance of his ideas any less remarkable.

¹BT, 71 / KSA 1, 97

For Nietzsche, Socrates and his followers are responsible for the demise of ancient Greek culture, grounded in an aesthetic conception of human life, and for the birth of modern European culture, driven by a new and unprecedented faith in reason, coherence and logic. Whereas ancient Greek wisdom was both constructive and destructive, grave and gratuitous, Apollonian and Dionysian, Socratic wisdom is no longer fuelled by this antagonism. It is the result of Apollo's triumph over Dionysos and the beginning of truth's historical and philosophical empire. In this new cultural environment, spontaneity and arbitrariness are no longer encouraged or admired. The logical element takes hold of imagination and metaphors are turned into concepts, precepts and laws. Reason is no longer a vehicle for happiness and delight, but a measuring instrument placed above human life.

Socrates is the first and most extreme exponent of this new cognitive attitude. He regards the attainment of knowledge as the highest of human conquests and spends his entire life pursuing it. In his eyes, truth and value are one and the same thing: whatever is to count as beautiful, pleasant or noble must be submitted to a logical examination and shown to be rational. What is more, life itself must be subjected to this exam: Socrates' main concern is no longer whether life is pleasant or captivating, but whether it makes sense. His philosophical ideal is a life free from contradiction and self-deception, a life where every choice is the result of a rational judgement.

11.1 The Socratic Turn

In Nietzsche's writings, the revolution ushered in by Socrates is often illustrated by the analogy between living and dreaming, or the idea that human existence amounts to a dream-like existence.² This image is also a well-known Platonic motif, and the coincidence is certainly not accidental: its philosophical implications are closely related, once again, to the 'problem of science', which unites and separates the Platonic and the Nietzschean conceptions of truth.

Ancient Greek life, under the aegis of Apollo, amounts to a long and blissful dream, where pain and suffering are redeemed by the transfiguring power of art, but the arrival of Socrates is a moment of awakening. Apollonian men and women revel in their own inventions, like dreamers who know they are dreaming and wish to go on dreaming, but Socrates refuses this self-indulgence. Turning his 'great Cyclopaean eye' on this fantastic dreamworld—an eye 'in which the lovely madness of artistic enthusiasm never glowed'³—he subjects its fabrications to a new and cruder kind of examination. With a few stern words, images turn into concepts, their truthfulness is challenged and their spell is finally broken: magic and myth lose their power, gods and nymphs disappear into thin air, the whole edifice of human invention comes under threat. Its charm is lost on rational individuals, to whom clarity and utility are the only concerns.

²See, for ex., BT, 25 f., 150 f. / KSA 1, 38 f., 886 f.; KSA 1, 760.

³BT, 67 / KSA 1, 92

At first sight, this destruction seems justified. Beauty was sacrificed, yes, but in the interest of truth and knowledge. A world of illusion was replaced by a more real world, where ideas and judgements are closer to the actual being of things. But as personal experience usually shows, what is peculiar about dreams is precisely the difficulty in distinguishing what is real from what is not, what belongs to the oneiric sphere and what does not. In most dreams, the dreamers are convinced they are awake, and only by actually waking up do they recognize their mistake. Dreams are unsettling because they mimic the feeling of awakesness that is typical of one's real, non-oneiric existence. And therefore, even when the dreamer feels he or she has awoken, this feeling may amount simply to a new and more sophisticated form of dreaming.⁴

In the initial sections of this book, when discussing Plato's philosophy, we came across a very similar phenomenon. We saw that his critique of the doxastic nature of human cognition does not amount to replacing a false account of reality with a true one. His discovery is a negative discovery, just like that of a dreamer who suddenly realizes he or she is dreaming. The dialogues highlight the fact that none of the ideas and words we resort to in daily life are as reliable or as intelligible as we take them to be. Their usual familiarity is an illusion whose spell, once it is recognized as such, loses its power. The experience brought out in Plato's works is therefore one of *awakening*. However—and herein lies the main point—it does not lead to an *actual* awakening, but merely to an awakening *to the idea that one is dreaming*. As soon as it becomes clear that one's account of reality is not real, but illusory, one is faced with a different version of the same oneiric world. Gods and nymphs, but also everyday objects and ideas are suddenly rendered opaque and awkward, like the props of a theatre play that have lost their dramatic purpose. In Plato's words, this unusual experience exposes the 'long day' of human existence as a 'nocturnal day', caught between sunlight and darkness.⁵

The world unveiled by Socrates is at best a 'middle world', located somewhere between waking and sleeping. It is the old familiar world devoid of its familiar meaning—a strange and unwelcoming world, where our usual confidence in the ability to interpret reality is suddenly compromised. We are left as strangers in a strange land, or as fish out of water, in need of the vital oxygen provided by dreams.⁶ And our main concern is thus to escape this confrontation and go back to sleep.

Despite the obvious parallel between Plato's and Nietzsche's images, there are at least two major differences between them. And to grasp their meaning is to grasp the main divergence between Plato's philosophical project and Nietzsche's metacognitive project.

⁴On the ambiguity between dreaming and being awake, see, for ex. BT, 151 / KSA 1, 887; BGE, 68, 82 f. / KSA 5, 97, 114 f.

⁵See *Re* 521c.

⁶See *Phd* 109a–110a

Firstly, although Plato's undertaking is a negative undertaking, he seems to think that it is nonetheless a necessary one. Or rather, although his critique is not a sufficient condition for the actual liberation from the dreamworld of ordinary reason, it is nonetheless a necessary condition for a *positive* or *post-critical* transformation ever to come about. This belief can be expressed with the help of a brief statement by Albert Camus, offered in a very different context but with the same idea in mind: 'to accept the absurdity of everything around us is one stage, a necessary experience: it should not become an impasse. It arouses a revolt that may become fruitful.'⁷ Translated into Platonic language, this means that the only way out of the doxastic realm is the 'purge' (καθαρμός) prescribed in the *Sophist* and the *Republic*.⁸ To be sure, exposing the spuriousness of human beings' cognitive self-confidence will not automatically provide them with actual knowledge. However, it will bring them closer to it. As Socrates points out, in a passage already quoted above, the awakening proposed by philosophy can lead to one of two results: 'either we shall find what we are after [namely, knowledge], or we shall be less inclined to think we know what we do not know at all; and surely even that would be a recompense not to be despised.'⁹ Those who are awoken by philosophy are better off either way, for they are *less deluded*, and hence *less imprisoned* than ordinary dreamers.

For Nietzsche, however, there is nothing to gain from this transformation. On the contrary. Since every human thought is grounded in a metaphor, human reason as such is intrinsically oneiric, and Plato's critique only succeeds in replacing one kind of dream with another. The formal parallel between the two positions is clear: just as Plato's Socrates sought to expose human beings' usual standpoint as a dream-like form of awokeness, Nietzsche aims to expose Socrates' own shrewdness as a dream-like form of awokeness. But his point is not simply that all of us are dreamers, *no matter what*. Although human cognition is indeed oneiric, not all dreams are alike. And while the first, pre-conceptual or pre-Socratic kind of dreaming is free and life-promoting, the second, conceptual or Socratic one, is slavish and debilitating. Whereas Socrates' oneiric awokeness leads to a fruitless standstill, in which life itself is called into question, Apollonian dreaming is smooth and fertile. It is not a rational phenomenon, but an artistic one, and its fruits are not conditional on a painful or laborious quest. Unlike the philosophical dreamer, the artistic dreamer hangs on to a ceaseless and ever-changing illusion, with no need or interest to let go.

'Leave him hanging', cries *art*. 'Wake him up', cries the philosopher, in the pathos of truth. But he himself, convinced that he is shaking the sleeper, falls into an even deeper magical sleep—and dreams perhaps of 'ideas' or immortality. Art is more powerful than knowledge, for the former wills life, while the latter achieves as its final goal only—destruction.¹⁰

This antithesis leads to a further point of contention between the Platonic and the Nietzschean projects, related not only with what Nietzsche termed the 'will to truth', but also with what he termed the 'will to illusion'. As both Plato and Nietzsche recognize, Socrates' standpoint is an exceptional standpoint. His quest for truth is

⁷Camus, *Essais*, 1425

⁸*So* 230b–c and *Re* 533c–d. See Sect. 5.2.

⁹*Th* 187c1–3

¹⁰KSA 1, 760

too radical, too exacting and too contrary to life's practical demands. Consequently, human beings are hardly ever awake, in the Socratic sense of the world. Their adherence to Socrates' ideas is always brief and reticent, like that of 'people awakened from a nap'.¹¹

Yet although both ancient and modern individuals are natural born dreamers, the latter's dreams are very different from the former's. Whereas Apollonian men and women regard life itself as a dream, and human cognition as a globally metaphorical process, modern dreams are spoiled by the earnestness of Socrates' quest. Ancient human beings do not long to be awoken because for them illusion is itself the truth. There is nothing beyond the world of dreams because the world itself is a dream: its truth is the beauty offered to the dreamer and its rules are the magic and incoherent rules of dreams.

Ancient men and women 'knew how to dream and didn't even need to fall asleep first'.¹² Modern men and women no longer possess this ability. After being awoken by Socrates and losing touch with life's metaphorical roots, they went back to sleep. But their sleep is heavier than that of ancient human beings: their dreams are no longer self-conscious, *creative* dreams, but involuntary, *conceptual* ones. Modern human beings are no longer aware of their dreaming condition. They mistake the world of dreams for a world in itself, whose meaning is a matter of truth, order and intelligibility. Their artistic powers are no longer free or spontaneous and illusion is no longer lived as illusion, but as the opposite of truth.

In this new, self-deluded existence, dreaming itself is only tolerated as a state of exception, as a kind of insanity opposed to the seriousness of concrete everyday life. According to Nietzsche, however, it is one of the few remaining bridges between our current and our former or original selves. Along with sudden outbursts of passion, bouts of madness, drunkenness or the ecstasy of artistic creation and enjoyment, dreams rank among the few experiences that are still capable of unweaving reason's conceptual web and tearing down the 'prison walls' of objectivity.¹³

The dialectic of dream and reality is another important element of Nietzsche's overall characterization of the relationship between life, truth and freedom. In particular, his reinterpretation of the Platonic dialectic of dream and reality is another means of challenging Socrates' scientific revolution and his own views on the relationship between life, truth and freedom. As the previous discussion has shown, the different degrees of awokeness distinguished by Nietzsche correspond to different degrees of cognitive freedom, that is, to different degrees of practical and intellectual autonomy. Accordingly, the oneiric imagery used by Nietzsche is akin to the carceral imagery employed, for instance, in Plato's allegory of the cave. Central to both cases is the assessment of human freedom 'as regards both education and the lack of it'.¹⁴ And the varying quality of human beings' outlook on reality translates, in both cases, into very different existential scenarios.

¹¹ *Ap* 31a4

¹² *GS*, 70 / *KSA* 3, 423

¹³ On the idea that dreams provide an insight into humankind's aesthetic/oneiric infancy, see, for ex., *HAH*, 16ff. / *KSA* 2, 31ff.; *D*, 157 / *KSA* 3, 226; *GS*, 70 f. / *KSA* 3, 423.

¹⁴ *Re* 514a2

In Plato's version, our usual standpoint is an imprisoned standpoint. However, it is also a sleeping standpoint, whose understanding of reality resembles that of a dreamer. Philosophy's role is to awake us and confront us with our imprisoned condition. This radical ἀναγνώρισις is precisely what takes place inside Plato's cave. Once the prisoners are allowed to turn their heads, they realize that their outlook on reality is seriously limited: With this realization, they are disabused of their former beliefs and left more awake than before and more capable of assessing their practical situation. But this awakening is not yet the end of human dreaming, nor the elimination of human captivity. It is merely the first, negative step of a wider philosophical revolution, whose intended result is an actual awakening, i.e. an actual departure from Plato's cave.

In Nietzsche's version, human consciousness is originally free. It is not free to choose or to decide, to interpret or to react to the surrounding world, but it is not yet doomed to think in terms of true and false concepts, of accurate and inaccurate versions of reality. It is also a dreaming standpoint, but its dreams are not opposed to reality as appearance is opposed to truth. With the rise of Socratism, this freedom is gradually compromised and consciousness is forced into the epistemic framework of modern reason. For Nietzsche, the awakening to the fact that one is unable to grasp reality *as it really is* does not represent a step towards the liberation of human consciousness, but the first stage of humankind's self-imprisonment. In his eyes, men and women are not originally imprisoned: they have *become* so due to an exaggerated and unhealthy fascination with truth. Socrates, as the main promoter of this new cognitive trend, is not the liberator, but the jailer of mankind; he has helped tie up human consciousness, only to lament afterwards its inability to break free.

Since this transformation is about freedom and autonomy, its effects are not limited to the domains of logic or epistemology. Nietzsche regards the rise of Socratism as the starting point of a philosophical movement with aesthetic, moral and political implications; a movement whose underlying principles have gained strength throughout the centuries and whose far-reaching influence is ultimately responsible for the cultural crisis of modern Europe.

11.2 Athens and Jerusalem

As is well known, Nietzsche paid special attention to what he considered to be the moral and religious implications of Socrates' and Plato's scientific revolution. In this new intellectual attitude he saw not only the origin of the modern distinction between a phenomenal and a noumenal world, but also the seed of a metaphysical tendency to endow truth with an eternal or transcendent value. Once reason was subjected to Socrates' impossible demands, its natural limitations came to the fore: human ideas were reduced to mere beliefs, aimed at capturing a mysterious and ever-fleeting world in itself; as a result, the notion of truth was detached from the alleged truthfulness of this or that human belief and placed in a higher and inaccessible plane. In like manner, once the value of human choices was made to depend on a rational design, life itself was soon found wanting. Human existence emerged as a

faulty, downgraded existence and its contradictions led to the positing of a truer or higher form of existence. This shift from a subjective *here* to an objective *beyond* represents, for Nietzsche, a form of self-imprisonment. More than that, it constitutes the basis of a slavish form of cognition, directly opposed to human beings' original sovereignty. Socratic men are not in charge of their thoughts or judgements, nor are they able to create their own values. Their entire world is sustained by the constant adaptation to external forms of valuation, and this dependence amounts to a form of weakness and servitude: it betrays the will to be 'ruled' or 'commanded', to be a creature rather than a creator.¹⁵

At the hands of Socrates and his followers, truth became not only an epistemic imperative, but also a moral one. The historical and conceptual link between logical truth and moral truth is a central Nietzschean theme, and the key to understanding Nietzsche's insistence on the relationship between Socrates, Plato and the Judaeo-Christian tradition. In the 1886 preface to the *Birth of Tragedy*, Nietzsche suggests that his first book, which offered his first approach to the 'problem of science', is already dedicated, albeit indirectly, to the issue of morality. More specifically, it already puts forth a philosophy set 'against the *moral* interpretation and significance of existence.' His aesthetic characterization of ancient Greek wisdom allows only for an 'artist's meaning . . . behind all that happens—a "god", if you will, but certainly only an utterly unscrupulous and amoral artist-god who frees (*löst*) himself from the dire pressure of fullness and *over-fullness*, from *suffering* the oppositions packed within him, and who wishes to become conscious of his autarchic power and constant delight and desire, whether he is building or destroying, whether acting benignly or malevolently.'¹⁶

With Socrates, however, this power is denied. Human autarchy is sacrificed to a new divinity, namely the notion of truth, and life acquires for the first time an overtly moral dimension. The parallel between truth and morality constitutes one of the main conceptual bridges between Nietzsche's early writings, primarily focused on the value of truth (and the possibility of an aesthetic redemption of human cognition) and his later writings, increasingly concerned with Jewish and Christian morality (and the possibility of a complete 'transvaluation' of moral values). Nietzsche's interpretation of the will to truth as a religious or ascetic symptom marks a shift of direction from Athens to Jerusalem, from his early Dionysian worldview to a fierce attack on Christianity.

As Nietzsche writes in *On the Genealogy of Morality*, 'there is no "presuppositionless" knowledge, the thought of such a thing is unthinkable, paralogical: a philosophy, a "faith" always has to be there first, for knowledge to win from it a direction, a meaning, a limit, a method, a right to exist.' Therefore, whoever, like Plato, 'tries to place philosophy "on a strictly scientific foundation", must first stand on its head not just philosophy, but also truth itself.'¹⁷ Although

¹⁵See KSA 10, 61; GS, 205 f. / KSA 3, 582f.

¹⁶BT, 8 / KSA 1, 17

¹⁷GM, 112 / KSA 5, 400

Plato's philosophical project aims to identify and suspend all of reason's dogmatic assumptions, it is itself supported by a dogmatic faith in the referential value of truth. In this respect, its formal structure is similar to that of many other moral or religious systems: by ranking every idea or judgement below the stage of science, Socratism takes the form of a negative theology, where the value of human beliefs is determined by their degree of imperfection. In this negative scheme, ideas and judgements are mere *imagines scientiae*, bound to emulate a truth they are unable to convey; and philosophy's task is to raise these images from the negative terrain of self-conceit to the neutral terrain of ignorance, and from then on to the positive terrain of science.

This elevation requires, therefore, that everything be tested with regard to its scientific worth. Reality can no longer be deemed beautiful, or pleasurable, or useful: it must be shown to be such. Contradiction and error are no longer allowed and experience is divested of its usual spontaneity. Wherever Socratism 'directs its probing gaze, it sees a lack of insight and the power of delusion, and it concludes from this lack that what exists is inwardly wrong and objectionable.'¹⁸ As a consequence, the phenomenal world is reduced to the more or less distant or distorted echo of a scientific or noumenal world, concealed behind the doxastic realm.

No doubt, those who are truthful in that audacious and ultimate sense which faith in science presupposes *thereby affirm another world* than that of life, nature, and history; and insofar as they affirm this 'other world', must they not by the same token deny its counterpart, this world, *our world*?... But you will have gathered what I am getting at, namely, that it is still a *metaphysical faith* upon which our faith in science rests—that even we knowers of today, we godless anti-metaphysicians, still take *our* fire, too, from the flame lit by the thousand-year old faith, the Christian faith which was also Plato's faith, that God is truth; that truth is divine...¹⁹

Nietzsche's engagement with Plato's philosophy is usually broad-scoped and rarely comes down to the actual argumentative detail of the dialogues. Plato and Platonism are often used as general historical labels, which tend to downplay the complexity and heterogeneity of the Platonic corpus. In this respect, Nietzsche's attitude is hardly an exception. Nevertheless, the previous passage should not be too quickly construed as an example of the 'two-world reading' usually associated with Plato.²⁰ Nietzsche's aim, here and elsewhere, is to show that a philosophical project such as Plato's, grounded in a wholesale criticism of phenomenal reality, is naturally open to this kind of interpretation. More than that, Nietzsche points out that Plato's philosophy has indeed been interpreted in this way, as it was de facto one of the main philosophical sources of Christian thought. And here, once again, his argument is not particularly original. Insofar as Christianity is based on a two-world cosmology, which echoes the fundamental contrast between an earthly and a divine standpoint, it does amount to a historical development of Platonism—a 'Platonism for the "people"',²¹ elevated to a world religion.

¹⁸ BT, 66 / KSA 1, 89

¹⁹ GS, 201 / KSA 3, 577

²⁰ See Sect. 4.1.

²¹ BGE, 4 / KSA 5, 12

But this dualism is not the only parallel between the Platonic and the Christian traditions. In the preface to *Beyond Good and Evil*, Nietzsche credits Plato not only with having invented ‘the pure spirit’ (*der reine Geist*), but also ‘the good in itself’ (*das Gute an sich*), thereby paving the way for ‘millennia’ of ‘Christian-ecclesiastic pressure.’²² Concerning this second claim, we have seen earlier, when discussing Plato’s diagnosis of human cognition, that the Platonic notion of good is primarily a *formal* notion, placed at the top of reason’s conceptual alphabet.²³ Since human beings are by nature self-interested beings, human life is continually oriented towards what is perceived as the greatest or highest good available. Whether in a conscious or unconscious manner, all human judgements and choices are endorsed in view of what is deemed, at each given moment, the best or more advantageous alternative or course of action. But although ‘the good’, in this generic sense, is indeed ‘the end of all our actions’,²⁴ it is still a formal concept, able to accommodate an infinity of different contents. Although every one of us aims at the greatest good, not all of us agree on what the greatest good actually is, and this disagreement is precisely what fuels many of Plato’s dialogues.

Socrates’ enquiries are often debates about what is *best*—namely, what is the best course of life, what is the best conduct in a given situation, what is the best kind of education, and so on. The dialogues revolve around whether it is best, for instance, to be just or unjust, to pursue philosophy or rhetoric, to embrace a life of pleasure or one of discipline, to take on a lover or to refrain from doing so, etc. In these disputes, Socrates’ usual role is that of a referee: his first aim is to show that the answers to these queries are always specific declensions of τὸ μέγιστον ἀγαθόν, that is, specific *versions* of what the greatest good, in these situations, actually is; his second aim is to show that these versions, when subjected to a closer examination, are usually problematic. When questioned about their reasons for endorsing this rather than that alternative, Socrates’ interlocutors are led to contradict themselves and to lose faith in their initial stances.

What emerges from this procedure is once again the need for a philosophical re-education of ordinary consciousness. Even though human life is naturally oriented towards the good, one’s idea of the good is usually misguided and self-contradictory. As a consequence, what is perceived as good or advantageous may turn out to be neither, and lead to evil or harmful consequences. For the good to be actually beneficial, it must be accounted for *scientifically*, that is, it must be *known* to be good, rather than merely *claimed* or *believed* to be so. In other words, the good must be elevated from the doxastic to the epistemic realm, from οἶσθαι εἰδέναι to actual εἰδέναι. And this elevation is philosophy’s most important task: the continual effort to correct the discrepancy between the version of good adopted by consciousness and good itself.

²² BGE, 4 / KSA 5, 12

²³ See Sect. 1.3.

²⁴ *Go* 499e8–9

For Nietzsche, ‘Plato wanted to use all his strength (the greatest strength a philosopher had ever had at his disposal!) to prove to himself that reason and the instincts converge independently on a single goal, on the Good, or “God”; and, ever since Plato, all theologians and philosophers have been on the same track.’²⁵ Yet insofar as the Platonic idea of good is a formal idea, it is still very far from the Christian idea of good and the more concrete imperatives of love, piety or charity. Indeed, τὸ ἀγαθόν is above all a riddle, lacking a definitive solution, and Platonic morality is likewise interrogative rather than affirmative or dogmatic. The transition from Plato to Christian theology is therefore not as straightforward as it might appear. For the Christian believer, the greatest good is no longer a question, but a certainty; although its true face may remain hidden, unfathomable, beyond human reach, it is nonetheless irreducible to a dialectical or philosophical demonstration. Whereas for Socrates and Plato intelligibility and coherence are basic conditions of truth, Christianity subordinates both of these requirements to the immediacy of religious faith. For the Christian, one must ‘believe in order to understand’²⁶; truth is not a rational or scientific conquest, but the content of a divine revelation, open to philosophers and non-philosophers alike.

If Plato’s idea of good is construed in a purely formal way, Nietzsche’s claim that ‘the Christian faith is also Plato’s faith’ is no better than a sweeping generalization. But this formal approach is not the only possible approach to Platonic morality, nor indeed the one usually favoured by Nietzsche. In the *Gorgias*, for example, as the different contenders try to reach an agreement concerning the best or most advantageous way of life, Socrates strives to show that their proposed solutions—namely, the life of rhetors, the life of despots or the life of hedonists—are not as good or as advantageous as they appear. Yet Socrates’ role is not confined to this negative procedure. Apart from his immanent critique of the different views set forth by his interlocutors, he also advances a doctrine of his own. After having argued, against Polus, that the despot is not necessarily powerful, he goes on to claim, quite vehemently, that to do evil is worse than to suffer it, and even worse when the evildoer is left unpunished²⁷; and after having argued, against Callicles, that the hedonist is not necessarily happy, he goes on to claim, with no less conviction, that the just man is the happiest of all men, and justice the greatest of all goods.²⁸ If Plato’s aim were purely elenctic, the dialogue would consist simply in a systematic *reductio ad absurdum* of the positions held by his opponents. In the end, the greatest good, or the best way of life, would remain a formal category, lacking a satisfactory or coherent content. The reader would be told what the best way of life *is not*—namely, rhetoric, despotism or hedonism—and encouraged to go on looking. Besides, if Plato’s intention were to expose the doxastic nature of human cognition in a more radical or complete manner—as he appears to do, for example, in the

²⁵ BGE, 81 / KSA 5, 112

²⁶ Isa 7:9. See Sect. 7.4, fn. 84.

²⁷ *Go* 468e–469b

²⁸ *Go* 527b–e

Theaetetus—his critique would have to move beyond despotism and hedonism, good and evil, and extend to the entire alphabet of human reason. He would then merely establish, in a more general manner, that human beings are ‘of no account in respect to wisdom’²⁹—either of what is best or, indeed, of *anything else*. In the *Gorgias*, however, Plato’s aim seems to be different. He is concerned not only with disproving his opponents’ immoral claims, but also with introducing a moral worldview that is both intelligible and feasible.³⁰

Plato’s departure from the strictly negative strategy usually associated with Socrates is also noticeable, for example, in the *Republic*, where the Platonic Socrates is equally categorical about the merits of a just and virtuous life, or in the dialogues’ eschatological myths, where the exaltation of virtue, the immortality of the soul and the idea of a Final Judgement play an important part. All of these texts denote a moralizing tendency that is arguably irreducible to the effort of ‘casting out the conceit of knowledge’³¹ underlying ordinary human cognition. However, none of these passages is a definitive proof of Plato’s moral dogmatism. Indeed, the very form of the dialogues should caution interpreters against such categorical conclusions—either because Plato’s supposed seriousness is often hard to distinguish from Socrates’ irony or because what appears to be final in his dialogues is often only provisional, open to new and more subtle forms of refutation.

In any case, it is fairly clear that Nietzsche’s Plato is not a purely dialectical or elenctic Plato, bent on turning every human belief against itself. Nietzsche’s critique of Platonic morality appears to be focused, rather, on the dogmatic attitude featured in

²⁹ *Ap* 23a

³⁰ The *Gorgias* ends with a practical exhortation that encapsulates its entire pedagogic programme: ‘Let us therefore take as our guide the doctrine now disclosed, which indicates to us that this way of life is best, to live and die in the practice alike of justice and of all other virtues. This then let us follow, and to this invite everyone else.’ (*Go* 527e2ff). Of course this exhortation is still merely formal, insofar as the concepts of justice and virtue are not simple or self-evident, but Plato seems to be appealing here to a kind of life that is nonetheless familiar to his readers, if hard to put into practice.

Socrates’ affirmative stance in the *Gorgias* may be due to pedagogical and political concerns. In his brief introduction to the dialogue (ed. cit., 249), W. Lamb points out that Plato’s ‘design of attracting the attention of the ordinary man of some culture is evident’. Alternatively, Socrates’ position can be interpreted along the lines of Pascal’s famous *pari*. That is to say, by refuting Polus’ doxastic defence of injustice and by favouring his own doxastic defence of justice, he may be choosing between the lesser of two evils. Whereas doing wrong amounts to openly embracing a mistaken standpoint, doing good amounts to embracing a standpoint that has not yet been refuted, and can still be proven right. Morally speaking, the first of these options is much more hazardous than the second one. . . . Finally, it might also be argued that the whole dialogue is written not from a doxastic standpoint, but from a philosophical or epistemic standpoint. If Socrates is already in possession of such a standpoint, he may be *assuming* the epistemic validity of his defence of justice without bothering to prove it. In this case, the dialogue’s aim would simply be to refute Polus’ and Callicles’ mistaken theses, while a proper demonstration of the correct alternative would be left for another occasion.

³¹ *So* 230b1

the *Republic*, the *Gorgias* or the *Phaedo*.³² His Plato is first and foremost an *edifying* Plato: the champion of justice and selflessness, of honesty and resignation. He is the main vehicle for the historical and cultural affirmation of a new and less ambiguous brand of dogmatism, to which Nietzsche will dedicate most of his later writings.

11.3 Masters and Slaves

As Nietzsche's main focus shifts from a critique of scientific reason to a critique of morality, his genealogical account of freedom is likewise transformed. In his early writings, freedom and slavery are chiefly associated with the contrast between an arbitrary or aesthetic mode of cognition and a logical or conceptual mode of cognition. In the former case, human beings were abler to create their own values, as artists creating a work of art; their life resembled a conscious dream, where logical and conceptual oppositions were somehow less binding than they are today. In the latter case, human life was confined to the 'prison walls' of science³³ and subordinated to an increasingly complex web of concepts, rules and regulations. Valuation became a passive rather than an active process and the real world was depreciated in favour of a new, ideal one, where truth reigns supreme.

To the later Nietzsche, humankind's primordial freedom is more directly linked to the notions of authority, strength and power. In *On the Genealogy of Morality*, for example, the problem of freedom is once again framed by an evolutionary account of human cognition. Primitive human beings are once more portrayed as naïve and capricious beings, incapable of the intricate calculations introduced by conceptual reason. They live above all for the present moment, moved by strong feelings rather than coherent or persuasive truths. Their value judgements 'are based on a powerful physicality, a blossoming, rich, even effervescent good health that includes the things needed to maintain it, war, adventure, hunting, dancing, jousting and everything else that contains strong, free, happy action.'³⁴ However, instead of emphasizing the playful or artistic aspects of this primitive existence, Nietzsche highlights the violence and the cruelty it necessarily entails. In this primitive world, life is a permanent struggle and every action is simultaneously a display of strength and domination. Reason is not yet distinguished from the drive to satisfy one's most basic instincts, to sate one's immediate needs and appetites. Therefore, prediction and calculation are still overshadowed by a primitive scorn for safety and comfort, by a 'shocking cheerfulness and depth of delight in all destruction, in all the debauches of victory and cruelty.'³⁵

³²In *Beyond Good and Evil*, Nietzsche contrasts Socrates' scepticism with Plato's moral naïvety. See BGE, 81 / KSA 5, 112: whereas Socrates 'had seen through to the irrationality of moral judgments', Plato 'was more innocent in such matters and lacked Socrates' plebeian craftiness.' His goal was to affirm the divine nature of the Good, thereby anticipating the task of 'all [future] theologians and philosophers.'

³³See BT, 148 / KSA 1, 883; KSA 1, 760.

³⁴GM, 17 / KSA 5, 266

³⁵GM, 23 / KSA 5, 275

Yet this savage innocence was short-lived. Just like the reckless artists celebrated in Nietzsche's previous writings, these wild beings soon found themselves imprisoned within the confines of social life and led to conform to different rules and customs. In contrast to their original freedom, which hinged precisely on their unaccountability, their new, 'cultured' existence demanded the limitation of their instincts. As their power encountered that of a community, individual strength and cruelty bowed to measure and reasonableness; the call of nature was bound by the shackles of coherence and peace, of what is expected and allowed. And this transformation led to an important change in the very physiology of these primitive individuals: instead of beings made for action, to whom every resolution is simultaneously spiritual and physical, men and women were forced to detach the element of thought from the element of action and to grant it a new independence. They were made to rely on concepts, causes and effects—in a word, on their 'consciousness', the most fragile and undeveloped of all their organs.

This transformation led to a painful self-division. Natural drives and instincts could no longer be vented freely, but their call was still as audible and pressing as before. The original appetite for power and domination, part of the very essence of physical life, was not suppressed, but merely held in check. And the free individuals became caged beings, torn between their own power and the power of others, between freedom and punishment; trapped in a hostile and reproving world, no longer shaped in their image. As a consequence, their instincts were internalized, forced to find new and secret forms of satisfaction; they were concealed, dematerialized and translated into spiritual phenomena. More importantly, though, this inward movement was accompanied by a feeling of inadequacy and self-repudiation: social censorship became self-censorship and freedom itself became a source of shame.

This fool, this prisoner consumed with longing and despair, became the inventor of 'bad conscience'. With it, however, the worst and most insidious illness was introduced, one from which mankind has not yet recovered; man's sickness of man, of himself: as the result of a forcible breach with his animal past, a simultaneous leap and fall into new situations and conditions of existence, a declaration of war against all the old instincts on which, up till then, his strength, pleasure and formidableness had been based.³⁶

Herein lies the primary source of the moral worldview that will come to dominate human reason and human life. By turning against their primitive nature, human beings were led to draw a contrast between animal life and an ethereal or spiritualized life, placed on the opposite end of the ontological scale. This impulse of self-denial paved the way for new and increasingly spiritualized notions, such as selflessness, asceticism and self-sacrifice. With time, it also laid the ground for the more sophisticated notions of a redeeming God and a transcendent or heavenly existence. For Nietzsche, these ideas are the main forces behind the growing moralization of the modern world. In his genealogical account, the emergence of bad conscience and all its derivatives sets the stage for the 'moral sickness' redefined and aggravated by Platonism, Judaism and Christianity.

³⁶GM, 57 / KSA 5, 323

Nietzsche's account is evocative of a key Hegelian motif, namely the *Phenomenology's* 'unhappy consciousness', equally torn between a natural and a transcendent calling.³⁷ Nietzsche's 'bad conscience' is also triggered by a denial of life's intrinsic value and the attempted flight to a spiritual beyond. Moreover, Nietzsche's account is also associated with the asceticism of Judaism and primitive Christianity. In Hegel's case, however, this self-division is both necessary and liberating. The self-imposed wretchedness of the unhappy consciousness is merely the preparation for a higher phenomenological stage, leading to a dialectical reconciliation of self-identity and self-alienation. For Nietzsche, on the other hand, the evolutionary turn from the element of strength and worldliness to that of humility and self-denial is a self-diminishing movement. It leads humankind away, rather than towards, the actualization of its full ontological potential.

Nietzsche's genealogical approach to morality amounts to yet another instantiation of the 'problem of valuation', already explored with regard to truth and science. Just as Nietzsche sought to look beyond truth's referential value, by asking about the actual value of the notions of truth and untruth, he is now set on looking beyond the moral presuppositions embedded in modern reason, by questioning the actual value of the notions of good and evil, right and wrong, justice and injustice. To do so, he resorts to physiology, etymology and natural history, in an attempt to free philosophy and psychology from the moral presuppositions they have come to incorporate and to unveil a new and wider perspective on the issue of morality. A 'physiology of morality', as Nietzsche understands it, is meant to expose the contingency of moral concepts and to retrace their natural origin. Under this new optic, what is deemed moral or immoral is interpreted in light of the concrete demands of life and history, of health and physical growth. Instead of asking whether something is good or bad in itself, the physiologist must determine whether it is favourable to the development of the human type, whether it is a sign of health or a symptom of weakness and degeneration, whether it strengthens the will to live and to affirm oneself or undermines one's confidence and creativity.

The need for this kind of enquiry follows from human beings' spiritual emancipation and the rise of bad conscience. This should not be taken to mean, however, that primitive men and women were purely physical or mechanical beings, devoid of spiritual life. The difference lies rather in their understanding of the relationship between life's physical and spiritual components. Nietzsche aims to show that bad conscience is not a necessary trait of self-consciousness, but the result of a specific kind of self-consciousness, namely a *pathological* kind of self-consciousness. As the mind was detached from the body and turned against it, spiritual life acquired an independence it did not possess. Instead of fuelling the body's natural strength and its immediate cry for self-fulfilment and self-affirmation, it channelled its activity to an internal, disembodied world, increasingly devoid of physical expression.

³⁷ On this parallel, see notably Greene, 'Hegel's "Unhappy Consciousness" and "Nietzsche's Slave Morality"'.

This separation took its toll, firstly, on the primitive notions of good and evil. In opposition to contemporary theories regarding the origin of morals, which located the primary source of good and evil in the notions of utility and custom,³⁸ Nietzsche argues for the *aristocratic* foundation of these notions. In the primitive world, he claims, good and evil were not prescriptive or utilitarian notions, but *descriptive* ones. The word ‘good’ was usually applied to ‘good men’, that is, to powerful or superior men, whose talent for domination was immediately recognized and celebrated. The epithets ‘bad’ or ‘evil’, on the contrary, were the negative counterparts of this primitive nobility, applied to weakness and impotence. Nietzsche’s point, however, is not simply that good was synonymous with power or strength: he argues, rather, that the concrete meaning of this qualification was unilaterally established by the stronger or nobler class—that is, by those who possessed the actual power to make others conform to their definition of good. In its primary form, then, the good was the immediate reflection of a dominant and autocratic worldview. Good things were the attributes of a strong and noble life, viz. the noble man’s belongings, his physical complexion, his concrete achievements, his specific preferences, and so on. If these classifications possessed a moral content, it consisted simply in the noble man’s appreciation of his own intrinsic ‘goodness’.

This noble class was necessarily faced with a weaker class, on which it was able to exercise its sovereignty. But although weaker individuals were forced to comply with the values dictated by stronger ones, their obedience was accompanied by an internalization of their own unfulfilled instincts. Their repressed power was the first root of bad conscience—the psychological ‘swamp’, as Nietzsche sometimes puts it, where the first seeds of morality were allowed to grow. As the power of the primitive masters increased, so did that of their slaves, but in a negative way: initially, their submission was a source of shame and self-loathing, prompting a condemnation of their bodily existence and the unhealthy emancipation of life’s spiritual element; afterwards, this rising asceticism gave birth to a secret resentment against nobility and domination, which grew in direct proportion to the hardships and humiliations these individuals were forced to endure. Their inner strength was fuelled by the decrease of their actual strength, and a new kind of power—an indirect, spiritual, subterranean power—started to take shape.

Once again, Nietzsche’s account displays important similarities with Hegel’s characterization of self-consciousness. In the *Phenomenology* and the *Encyclopaedia*, primitive subjectivity is also defined in terms of power, strength and recognition. Although Hegel’s master-slave dialectic is part of a different philosophical project, with different aims and a very specific methodology, it can also be construed as an alternative approach to the ‘problem of valuation’. In Hegel’s account, the most basic form of self-consciousness consists in the recognition of another self-consciousness and its own drive for recognition. Initially, each self-consciousness is bound to look for itself outside of itself and to try and reduce the other to its own

³⁸The authors Nietzsche has in mind in the opening sections of *On the Genealogy of Morality* are Paul Ree and Herbert Spencer.

self. In this symmetrical struggle, subjectivity starts out as a pure will to domination, whose *terminus ad quem* is the complete suppression of otherness.

This confrontation takes the form of a mortal struggle, where each recognizing self seeks to eliminate the other. But since the suppression of one of the contenders cancels his ability to recognize the other, and thus the latter's very source of identity, this primitive scenario is replaced by a more complex dialectical arrangement. Instead of life and death, the struggle becomes a matter of authority and obedience, and the vanquished party is spared but reduced to slavery. Henceforth, master and slave are the protagonists of a new dialectical cycle, whose formal structure anticipates, to some extent, Nietzsche's account of the emergence of bad conscience. At first, the master is free to reap the fruits of victory, while the slave is bound to a life of servitude and dispossession. However, the slave's will to domination has not disappeared; it was merely thwarted by the master's triumph and forced to turn inwards. Gradually, nourished by work and self-discipline, it will reemerge as a new form of independence, enabling the slave to supersede both his own enslavement and the master's sovereignty.

Again, Hegel regards this whole process as logically necessary. Just as the master's sovereignty, due to its inner contradictions, is bound to be short-lived, so too is the slave's submission merely a stopping point in the path towards Reason and Spirit. Mastery, slavery and all the other phenomenological stages are necessary evils, as it were, without which human consciousness could not achieve its highest truth. In Nietzsche's account, on the contrary, the master's immediate power is qualitatively superior to the slave's internalized power. The replacement of the former with the latter represents, therefore, a qualitative loss.

A closer look at Nietzsche's basic argument can help clarify this difference. Once humans learned to sublimate their natural drives, turning the struggle against nature and each other into a struggle against themselves, self-consciousness was torn in two. Its physical half was neglected in favour of its spiritual half and the world became the outer setting of an inner and more essential form of reality. The pain and privation derived from weakness were imputed to the concrete side of existence and transformed into arguments against life and domination. As a result, the *natural* superiority of power, strength, cruelty and selfishness was replaced by the *spiritual* superiority of impotence, frailty, compassion and selflessness. This inversion of values was morality's first conceptual victory. Although stronger individuals did not need a moral worldview—their 'morality' was immediately determined by the concrete demands of health and vitality—weaker individuals were not as privileged. To make up for their natural inferiority, they were led to question the values of the strong and to transfer them to a new and unnatural region. The primitive, aristocratic conception of good was turned into its opposite and exposed as *evil*, i.e. as morally reproachable; conversely, the aristocratic conception of evil was turned into its opposite and revealed as *good*, i.e. as morally commendable. Moral precepts were born, therefore, from the artificial distinction between a 'master morality' (*Herrenmoral*), associated with humankind's primitive nature, and a 'slave morality' (*Sklavenmoral*), born out of an internalized will to domination.

As Deleuze pointed out, this line of argument is also central to Nietzsche's criticism of Hegel. Insofar as the latter's account of the master-slave scenario is not viewed from the *active* standpoint of the master, but from the *reactive* standpoint of the slave, it also denotes the slavish mentality Nietzsche so vehemently criticizes. Instead of the pure joy of domination that is characteristic of true masters, the ultimate aim of Hegel's contenders is 'to have their power *recognized*, to *represent* their power.'³⁹ But for Nietzsche this conception of power is entirely misguided. It is not the conception of the master, but that of the slave, 'the image that the resentful man has of power.'⁴⁰ And what is more, this passive attitude is the defining trait of dialectics in general, where negativity also plays a central role. Unlike the joyous immediacy of the master, the dialectician sets about negating whatever is deemed immediate and positive, only to derive from such negation a new and indirect form of self-affirmation.⁴¹

Whether criticizing morality in a stricter sense, in connection with the Judaeo-Christian tradition, or in the history of philosophy in general, Nietzsche's usual aim is to denounce its fundamental hypocrisy. More specifically, he aims to expose the fundamental contradiction between morality's self-appointed meaning and its *natural* or *pre-moral* meaning. Although the first advocates of morality exalted the merits of humility and selflessness, these very concepts were born out of greed and selfishness. Although they drew away for the physical world, they did so out of an intense but frustrated desire to be part of it. 'Slave moralists' were the inventors of a new and undeclared selfishness, which would prove far more deleterious than the innocent selfishness of primitive masters. While 'the noble method of valuation' 'acts and grows spontaneously, seeking out its opposite only so that it can say "yes" to itself even more thankfully and exultantly', the 'base method of valuation' is grounded in 'the *ressentiment* of those beings who, denied the proper response of action, compensate for it only with imaginary revenge. Whereas all noble morality grows out of a triumphant saying "yes" to itself, slave morality says "no" on principle to everything that is "outside", "other", "non-self": and this "no" is its creative deed.'⁴²

Because of the slave's calculating nature, born out of the experience of pain and humiliation, his power ends up surpassing that of the master. For the latter, cleverness is 'nowhere near as important as the complete certainty of function of the governing *unconscious* instincts, nor indeed as important as a certain lack of cleverness (*Unklugheit*), such as a daring charge at danger or at the enemy.'⁴³ Nobility is not about reason, but about instinct, spontaneity, audacity. For the

³⁹ Deleuze, *Nietzsche et la Philosophie*, 11

⁴⁰ Ibid.

⁴¹ From a Hegelian perspective, however, Nietzsche's defence of the master's standpoint amounts to an immediate, and hence abstract position, whose cogency can only be tested through the kind of enquiry Hegel proposes. See Houlgate, *Hegel, Nietzsche and the Criticism of Metaphysics*, esp. 7f.

⁴² GM, 20 / KSA 5, 271

⁴³ GM, 21 / KSA 5, 273

slave, on the contrary, power hinges on prudence and patience, on the willingness to wait and compromise. From a historical or evolutionary perspective, the Nietzschean contrast between the immediate or uncultured power of masters and the spiritual or educated power of slaves evokes the general educational path followed by primitive mankind: the evolution leading from natural ignorance (ἀπαιδευσία) to *political* education (primitive παιδεία), *technical* education (παιδεία τῶν τεχνιτῶν), and finally to the philosophical revolution advocated by Socrates and Plato. In the introduction to this book, we have seen that each of these leaps entailed the conviction of an increase in power, and therefore also in personal freedom. What Nietzsche points out, however, is precisely the opposite, namely that education is liable to bring about a loss of human beings' original freedom and the obligation to conform to dogmatic standards of valuation. Although Nietzsche's historical characterization of the transition from primitive mastery to primitive slavery is wide, heterogeneous, and thus not easy to locate within the educational progression I have sketched earlier, his main concern is to expose the *debilitating* rather than *empowering* effects of the previous forms of paideia. Primitive sovereignty, he argues, 'is the point where feeling reaches the opposite of the low temperatures needed for any calculation of prudence [*Klugheit*] or reckoning of usefulness—and not just for once, for one exceptional moment, but permanently.'⁴⁴ In other words, primitive sovereignty is not grounded in technical or scientific achievements: its power, as well as the freedom it provides, are due to its independence from every rational or moral rule.

For this very reason, however, primitive masters are simultaneously the most fragile of beings. Their childlike inability to plan and calculate, to sacrifice their immediate goals to more distant or abstract rewards, renders them particularly vulnerable to the world's arbitrariness. Their life is usually more pleasurable and intense than a life of submission, but it is also shorter and more perilous. They fight violence with violence, cruelty with cruelty, and shun every form of compromise. To borrow one of Nietzsche's favourite metaphors, these noble individuals resemble those rare and exuberant plants born in extreme natural environments and incapable of surviving in milder conditions. Weaker individuals, on the other hand, are amphibious and agile, solicitous and malleable. In this respect, Nietzsche's characterization of sovereignty is indeed similar to Hegel's, differing only in his appreciation of it. In the *Phenomenology*, masters are also doomed to extinction, but in favour of a superior form of self-consciousness; for Nietzsche, this extinction amounts to a degeneration—the movement whereby life is forced into the dogmatic moulds of science and morality.

⁴⁴GM, 11f. / KSA 5, 259

11.4 Asceticism and Nihilism

While Nietzsche's earlier contrast between aesthetic freedom and epistemic bondage was chiefly associated with the transition from the ancient Greek world to Socratism (and Christianity), the opposition between a master and a slave morality is granted a wider historical background. The first kind of morality is generally linked to a pagan or 'barbaric' world, represented by the cruelty of the primitive Germanic peoples, by the nobility of Homeric heroes or by the strength of the Roman generals. The second kind of morality is linked to the Jewish and Christian traditions. Judaism, in particular, represents for Nietzsche the slavish religion *par excellence*, the source of a vengeful hatred against all things noble and of the 'selflessness' that will come to dominate the Christian world.

This long genealogical progression is grounded in the basic fact of human suffering. For primitive masters, pain and violence were necessary and stimulating; they helped fuel and intensify 'all the debauches of victory and cruelty'. For weaker, unhealthier individuals, they were the cause of fear and unhappiness. Less equipped to protect themselves, they were also less able to satisfy their own craving for victory and satisfaction, which led them to turn inwards and resent the outer world. Starting from this basic diagnosis, Nietzsche recasts the history of human culture as a history of human impotence: he regards the moral, philosophical and spiritual transformations undergone by primitive, medieval and modern men and women as increasingly refined developments of the same defensive strategy; on his view, these transformations amount, for the most part, to the different strategies devised by powerless individuals to escape suffering and to find the satisfaction they crave.

By turning their back on the world, these slavish individuals came to embrace a series of 'ascetic ideals' capable of assuming different forms. The first one, already mentioned, is the feeling of shame or self-loathing that leads weak human beings to repress and revile their most basic instincts. Culturally, this idea is linked to the habits of fasting, abstinence and self-sacrifice, dating back to the very dawn of human civilization. Secondly, this feeling is extended to the whole world and to everyone else: the condemnation of one's actions and desires becomes a condemnation of acting and desiring as such. For the first time, life itself is put on the balance and found wanting; natural wisdom gives way to the 'wisdom of Silenus' and to a wholly new perspective on human existence. As a result, men and women retreat into their 'consciousness', their 'soul', their 'inner citadel', and become hermits, monks, Stoics. Thirdly, the world's former protagonism is projected into a new, spiritual world, along with human desires and aspirations. With this new inversion, the earthly element wins back its value: the physical world becomes a *sanctified* world, the product of a *divine* will. Since every earthly reality is an image or a sign of God, the hermit and the Stoic are no longer forced into isolation. The atonement for their weakness is now dependent on their moral conduct, on the selflessness required to earn a place in God's heavenly kingdom.

The idea of divinity, open to various meanings and interpretations, is the object of a series of important transformations. In the Greek world, the gods are not yet endowed with moral or saintly qualities, nor are they the negative image of human frustrations. They are the ‘reflections of noble and proud men in whom the *animal* in man felt deified.’⁴⁵ The Greeks ‘used their gods expressly to keep “bad conscience” at bay so that they could carry on enjoying their freedom of soul.’⁴⁶ With the Jewish and Christian traditions, all of this changes. The Jewish God is the avenger of human weakness, the source of power that the slave was unable to provide. God is the ultimate master, but also a reminder of human beings’ fallen condition, of their intrinsic wretchedness. With his negative conception of divinity, the Jew ‘seizes upon the ultimate antithesis he can find to his actual and irredeemable animal instincts, he reinterprets these very instincts as a debt [or guilt, *Schuld*] before God.’ In doing so, ‘he emits every “no” which he says to himself, nature, naturalness and the reality of his being as a “yes”, as existing, living, real, as God, as the holiness of God, as God-the-Judge, as God-the-Hangman, as the beyond, as eternity, as torture without end.’⁴⁷

The Jewish faith is the direct antithesis of the freedom enjoyed by Nietzsche’s superior beings. Its strength derives from the condemnation of all the things that made great individuals great: the violence and the selfishness that enabled them to flourish and assert their power are deemed immoral and replaced by a new and unnatural fixation on stillness and renunciation; the overflowing health they enjoyed is countered by a spiritual indifference to the human body, its needs and appetites; the Olympian carelessness of their deeds and demeanour is replaced by a morbid obsession with guilt, debt and accountability. All of these contrasts attest once more to the radical incompatibility between the aristocratic mode of valuation praised by Nietzsche and the slavish mode of valuation favoured by conventional morality. The former aims to lead human beings upwards, onwards, towards the fulfilment of their ‘highest potential power and splendour.’⁴⁸ The latter seeks to undermine their natural strength and vitality.

With the rise of Christianity, this contrast is aggravated. The emphasis is no longer placed on revenge and punishment, but on love, forgiveness and compassion. Life is no longer cursed, but sanctified, and the natural world is no longer shunned or reviled, but praised as a divine gift. For Nietzsche, this change of focus does not amount to a real transformation of the Jewish standpoint. Even though the world is accorded a new importance, its value is still entirely negative. Like the Jews, the Christians also say ‘no’ to ‘nature and naturalness’, but they do so in a more

⁴⁵GM, 64 / KSA 5, 333

⁴⁶GM, 64f. / KSA 5, 333

⁴⁷GM, 64 / KSA 5, 333

⁴⁸GM, 8 / KSA 5, 253

sophisticated way. From their perspective, earthly life is not meaningful or desirable in itself, but only as a vehicle for ‘redemption’, ‘salvation’, or ‘absolution’, notions whose meaning is dependent on the acknowledgement of the fallen or sinful condition of humankind. Whereas Judaism denounces earthly strength and earthly power as spurious and shameful achievements, Christianity calls for an actual celebration of weakness and compassion. In order to ‘redeem’ themselves, sinners must not only recognize their natural helplessness, but also elevate it to a moral imperative. Since salvation is primarily awarded to the poor-spirited, the humble and the destitute, true Christians must strive to become all of these things: they must turn their natural slavishness into a *voluntary* slavishness. Not surprisingly, this attitude represents, for Nietzsche, the most dangerous attack on human life. ‘By taking the side of everything weak, base, failed, [Christianity] has made an ideal out of whatever *contradicts* the preservation instincts of a strong life.’⁴⁹ Accordingly, the Christian God is but a ‘hybrid creature of ruin [*Verfalls-Gebilde*], made from nullity, concept, and contradiction, who sanctions all the instincts of decadence, all the cowardices and exhaustions of the soul!’⁵⁰

While Judaism is rigid and sectarian, Christianity is merciful and cosmopolitan. Its moral worldview is not the prerogative of a chosen people, but an appeal directed to humankind as a whole. In his later writings, Nietzsche charts Christianity’s historical legacy and is often engaged in detecting traces of the Christian outlook in different philosophical traditions. He sometimes praises the occasional resurgence of the aristocratic values embodied by primitive masters—during the Italian Renaissance, ‘a brilliant, uncanny reawakening of . . . the noble mode of valuation’⁵¹; or under the sway of Napoleon, in whom ‘the problem of the *noble ideal itself* was made flesh’⁵²—but is invariably led to recognize its exceptional nature. Modern European history is the history of a downfall, a domestication, a cooling down of humankind’s primal instincts. The raw innocence of warriors and aristocrats was overpowered by the spiritual cunning of priests, prophets and zealots. The words ‘God’, ‘sin’ and ‘salvation’ were translated into every language and absorbed by every philosophical system. Morality cast its web across the ages and left its sting in different regions and schools of thought. Everywhere the noble element was forced into new and hybrid mixtures, often at the price of extinction.

This declining tendency is well visible, Nietzsche argues, in the ‘theologian element’ that pervades modern German philosophy. ‘Kant’s success’, like that of Luther or Leibniz, ‘is a theologian’s success.’⁵³ As regards Kantian epistemology,

⁴⁹ A, 5 / KSA 6, 171

⁵⁰ A, 16 / KSA 5, 185

⁵¹ GM, 33 / KSA 5, 287. See also A, 5, 64f., 140 / KSA 6, 171, 250f., 359.

⁵² GM, 33 / KSA 5, 288. See also BGE, 87 / KSA 5, 120; KSA 12, 357.

⁵³ A, 9 / KSA 6, 177. Translation modified. See also, for ex., D, 2ff., 127ff. / KSA 3, 13ff.; 185ff.; GM, 115f. / KSA 5, 404f.; A, 9 f., 64f., 140 / KSA 6, 177f., 250ff.; 359f.

the distinction between a world of appearance and a world in itself, placed beyond the reach of human experience, represents for Nietzsche the new clothing of a much older religious ideal. In his eyes, all forms of idealism, by dislocating the essence of human life from a concrete here to an abstract beyond, amount to variations of the same theological deception. From a physiological point of view, their basic motivation is a feeling of weakness and their main result a slavish existence. Likewise, as regards Kantian morality, the ideas of virtue, duty, selflessness or a 'good in itself' are all part of morality's decadent vocabulary, favoured by weak and ascetic men. Here, again, the philosopher is at odds with the physiologist, and an obstacle to the attainment of human beings' 'highest potential power.'

Another important name in this historical survey is Schopenhauer, in whom the later Nietzsche sees one of the greatest and most dangerous enemies of the aristocratic ideal. Seduced by the Christian instincts of pity and compassion, by the Buddhist emphasis on silence and self-denial, Schopenhauer is the leading modern pessimist. His enthronement of the will is really an enthronement of the *negation* of the will⁵⁴: a condemnation of the 'penal servitude of volition',⁵⁵ revealing human beings' natural attachment to the world as illusory and burdensome. Hostile to life, to the senses, to the tumultuous joy of health and virility, Schopenhauer put life itself on the balance and concluded against its intrinsic value. His philosophy is one of the main sources of the 'will to nothingness' or the 'will to the desert' which Nietzsche considers the defining trait of modern European culture.

Finally, the aristocratic mode of valuation was further compromised by the modern dissemination of democratic and socialist ideas, whose first main surge dates back to the French Revolution. Nietzsche's reaction to this phenomenon is not hard to anticipate: in direct contrast to his *Herrenmoral*, reserved to masters and warriors, the democratic ideal advocates what he often describes as a *Herdenmoral*, a 'moral of the herd'. Broadly construed, democracy amounts to a re-enactment, in the political sphere, of the Christian ideal of redemption, directed to the weak and disenfranchised. Like Christian morality, this new attitude leads to 'a levelling and mediocritization of man',⁵⁶ to the 'degeneration and diminution of human being into the perfect herd animal', into 'a dwarf animal of equal rights and claims.'⁵⁷ Moreover, like Christian morality, democracy compels the strong and powerful to restrain their natural might and settle for a subdued existence. Its force derives from the social compact decried by Thrasymachus in Plato's *Republic*, forged by the weak as protection against the strong.

⁵⁴ See BGE, 45, 50f. / KSA 5, 68, 74f.; A, 149 / KSA 6, 173f.

⁵⁵ GM, 75 / KSA 5, 348.

⁵⁶ BGE, 134 / KSA 5, 183

⁵⁷ BGE, 92 / KSA 5, 127f. Translation modified.

These cultural and historical forces lead human beings away from the fulfilment of their true potential, towards a life of comfort, consensus and reasonableness. As the 'instinct of life' dies down, so do the conflicts and hierarchies that fuelled the ancient world. Morality, asceticism and democracy all lead to the triumph of weakness and exhaustion, to the nihilistic tendency that dominates the modern world:

Today we see nothing that wants to expand, we suspect that things will just continue to decline, getting thinner, better-natured, cleverer, more comfortable, more mediocre, more indifferent, more Chinese, more Christian—no doubt about it, man is getting 'better' all the time. . . Right here is where the destiny of Europe lies—in losing our fear of man we have also lost our love for him, our respect for him, our hope in him and even our will to be man. The sight of man now makes us tired—what is nihilism today if it is not *that*?⁵⁸

⁵⁸ GM, 57 / KSA 5, 323. As with pessimism, and despite Nietzsche's more frequent definition of nihilism as a sign of exhaustion, degeneration, weakness, and the like, he also distinguishes between a *positive, active* or *strong* nihilism and a *negative, passive* or *tired* nihilism. At stake in both cases is the devaluation of previously accepted values; but whereas the first kind of nihilism can be 'a sign of strength', or a call for the endorsement of new and stronger values, the second kind is 'a sign of weakness' or decadence, associated with Christianity or Buddhism. See KSA 12, 350ff., 367.

Chapter 12

Philosophy Unchained



Abstract Having analysed Nietzsche's critique of the cognitive imprisonment promoted by modern reason, I turn to his proposed solution, based on the replacement of a philosophy of truth or spirit with a philosophy of life, power and vitality. In Sect. 12.1, I discuss the possibility of a 'will to power', which is arguably the most important, but also the most controversial of Nietzsche's ideas. I argue that the ambiguity surrounding this idea is directly related to the unresolved tension already identified, in the previous chapters, with regard to the 'will to truth', characterized by Nietzsche both as an innate human feature and as a cultural and philosophical trend. In Sect. 12.2, drawing on this ambiguity, I claim that Nietzsche's *Machtphilosophie* can be read in light of Hegel's master-slave dialectic and of the debate in Plato's *Gorgias* about the relative merits of truth and power. More specifically, I argue that Nietzsche's apparent endorsement of the master's standpoint leaves him vulnerable to some of the criticisms raised by Plato and Hegel.

Keywords Nietzsche · Freedom · Philosophy · Power · Life · Socrates · Callicles

Regarding the issue of freedom, Nietzsche's philosophical project can be divided into two main moments, corresponding to a descriptive and a prescriptive approach to the 'problem of valuation'. In the previous chapters, we have dealt mainly with the first of these moments: namely, with Nietzsche's critique of human beings' usual mode of cognition. In this context, his critical strategy often consists in a) exposing our thoughts and actions as the products of a global system of valuation, open to different hierarchical configurations; and b) arguing that the mode of valuation we usually rely upon is dogmatic and misguided; it is not enabling of, but an obstacle to, the full actualization of freedom.

Nietzsche's criticism is centred on two complementary themes: he argues, firstly, that the referential value usually ascribed to the *epistemic* alternative between truth and untruth—whether inexplicitly, in everyday life, or explicitly, in philosophical and scientific debates—betrays an error in valuation, which must be questioned and corrected. As he sometimes puts it, our usual concern with truth speaks to our 'lack

of perspective' and the need for a wider and freer outlook on reality. Secondly, Nietzsche argues that the referential value usually ascribed to the *moral* alternative between good and evil—again both inexplicitly, in everyday life, and explicitly, in religious and philosophical contexts—is the product of a misevaluation, which must also be corrected. The 'will to truth' and the 'will to the good' constitute arbitrary modes of valuation automatically regarded as necessary and self-evident. They are the standards with which all things are measured, but their own value is hardly ever called into question.

Furthermore, both of these standards are incompatible with what Nietzsche takes to be the real driving force behind human life: a pre-epistemic and pre-moral (or *trans*-epistemic and *trans*-moral) *will to power*, to domination, to victory and self-enhancement. Whenever human existence is subordinated to epistemic and moral imperatives, the resources needed to fuel and intensify this primordial will are unduly restricted. Consciousness is denied the freedom to create its own values, or to change them at will, and the different possibilities of valuation are reduced to a more or less narrow set of fixed ideas, leading to a decrease rather than an increase in freedom and vitality. What is more, whenever the concern with truth and the good is elevated to a dogmatic creed, science and morality enter into direct contradiction with life: it is the case of asceticism and nihilism, promoted in different degrees by Judaism, Christianity and modern philosophy.

This contradiction brings us to the second, *prescriptive* moment of Nietzsche's project. Apart from denouncing the self-limiting or slavish nature of our usual standpoint, Nietzsche calls for a 'transvaluation of all values', starting with the ideals of truth and morality. He aims to transform the hierarchical structure of human values, so as to eliminate the artificial precedence of human cognition over human life, and it is this transformation we must now focus on, in order to grasp its exact nature, bring out its full implications and thereby judge the actual possibility of Nietzsche's proposed liberation.

12.1 The Will to Power

Throughout the previous sections, Nietzsche's reform of human cognition was generally associated with the idea of *power*. Instead of a life measured in terms of truth and untruth, right and wrong, good and evil, he advocates a kind of life where personal choices are endorsed or rejected depending on whether and how far they contribute to an actual increase of personal power. This basic formula, however, is open to different interpretations: 'power' (*Macht*), 'powerfulness' (*Mächtigkeit*) and the 'will to power' (*der Wille zur Macht*) are not self-evident notions; moreover, given the central role they play in Nietzsche's writings, determining their exact meaning is crucial to grasping the exact nature of the global transformation he advocates.

In the opening pages of *The Antichrist*, Nietzsche offers an explicit description of his new 'table of values':

What is good?—Everything that enhances people's feeling of power, will to power, power itself.

What is bad?—Everything stemming from weakness.

What is happiness?—The feeling that power is *growing*, that some resistance has been overcome.

Not contentedness, but more power; *not* peace at any price [*überhaupt*], but war; *not* virtue, but prowess [*Tüchtigkeit*].¹

In this description, as in others of the same nature, the new way of life proposed by Nietzsche is not simply about power, or the 'will to power', but also about 'growth', 'war' or 'prowess'. These different concepts are employed to qualify a general phenomenon that is associated, in Nietzsche's writings, with many other different terms and expressions, some of which used almost interchangeably, namely 'life' (or 'will to life', 'vitality', 'health', etc.), 'strength' (or 'will to strength', 'might', 'vigour', etc.), 'self-affirmation' (or 'joy', 'egoism', 'pride', etc.), 'self-enhancement' (or 'self-elevation', 'self-overcoming', etc.), 'domination' (or 'mastery', 'sovereignty', 'nobility', etc.), and the list could go on. All of these expressions refer to an existential regime whose exact contours are not entirely clear, not only because Nietzsche resorts to such a wide-ranging vocabulary, but also because he chooses to highlight, in different occasions, different aspects of this new mode of being.

Generally speaking, Nietzsche's alternative to the standpoints of truth, science, morality and all other dogmatic or metaphysical modes of valuation is a philosophy of power and domination; a philosophy that replaces the intellectual and moral deliberations typical of other systems of valuation with the *direct* experience of power and domination. But what exactly does this entail? And how does it differ from, say, a peculiar form of utilitarianism, or evolutionism, or hedonism? In order to make out, as clearly as possible, the meaning and scope of Nietzsche's standpoint, it is important to avoid a few hermeneutical pitfalls that he himself calls attention to in his writings.

To begin with, a mode of valuation that has moved beyond the dichotomies true/untrue and good/evil can be construed as a purely instrumental or Machiavellian mode of valuation: that is, as the outlook of a wholly selfish human being, entirely focused on his or her own self-interest. In this case, valuation would hinge solely on the maximization of personal advantage (whether in the form of power, gain, utility, etc.). Depending on the context, truthfulness and compassion would be just as valuable, if not more so, than deceit and cruelty, for both attitudes might yield useful or empowering results. But although this line of interpretation seems to agree with Nietzsche's insistence on the 'selfish' and 'ruthless' demeanour of superior beings, it contradicts his general definition of nobility. As seen earlier, utilitarian concerns are largely unknown to, and unworthy of, a noble consciousness: indeed, they represent, for Nietzsche, a form of enslavement.² Noble individuals are innocent and sincere, hot-blooded and childish, wasteful and exuberant; their joys are instantaneous and

¹ A, 4 / KSA 6, 170

² See KSA 11, 74.

not the result of calculation, nor the reward in view of which something else is pursued or undertaken. Planning and scheming are left to the weak and vengeful, to logicians and moralists, capable of patience, resignation and the laborious deduction of causes and effects. The noble individual lives for the present moment—which is why Nietzsche is sometimes careful to distinguish between power in general and the *actual feeling* or *sensation* of power, i.e. the immediate feeling that ‘power is *growing*, that some resistance has been overcome.’³ It is power in this concrete, physiological sense that he places at the heart of human existence.

Alternatively, Nietzsche also speaks of the ‘instinct of conservation’ and of an in-bred drive to protect oneself against destruction and death. He distinguishes between modes of behaviour that favour the prolongation of the human species and others that lead to degeneration and extinction. Often enough, things are deemed true, right or good that are unhealthy and pose a threat to human life; conversely, many of the things usually considered untrue, wrong or evil are in fact beneficial and life-enhancing. In light of these ideas, Nietzsche’s philosophy of power might be construed as a philosophy of self-preservation, where values are measured according to how far they can help extend and protect human life. In *Beyond Good and Evil*, however, Nietzsche explicitly rejects this line of interpretation, which he includes in a list of ‘philosophical prejudices’ to be overcome.⁴ Against Spinoza, Darwin and modern physiologists, he argues that the instinct of conservation is not life’s main driving force, but merely an indirect consequence of the will to power. This same point is also made in *The Gay Science*, where Nietzsche claims that the ‘the wish to preserve oneself’ is ‘a limitation of the truly basic life-instinct, which aims at the *expansion of power* and in so doing often risks and sacrifices self-preservation.’⁵ Furthermore, he also speaks of a noble ‘unconcern and scorn for safety, body, life, comfort’,⁶ as well of the joys of danger and the honour of going down fighting, as a hero in a battlefield.⁷

A third possibility is to interpret the ‘feeling of augmented power’ pursued by superior beings as a peculiar form of pleasure or gratification, and the ‘will to power’ itself as a peculiar form of hedonism. To be sure, Nietzsche does not encourage the avoidance of pain or conflict, nor does he prescribe a life of comfort and ease. These are the very ideals he attacks in his critique of morality, democracy and nihilism. On the contrary, he often speaks of hardness, discipline and the endurance of pain as signs of strength and nobility. However, these adversities are not valued out of a sense of duty or abnegation. Pain and strife provide the resistance which, when overcome, rekindles the feeling of power. They act as tonics, stimulating agents, whose intensity determines the intensity of the pleasure they are able to procure.

³A, 4 / KSA 6, 170. See also BGE, 122 / KSA 5, 167.

⁴BGE, 15 / KSA 5, 27f.

⁵See GS, 207f. / KSA 3, 27f., but also KSA 12, 260ff.

⁶GM, 23 / KSA 5, 275

⁷See, for ex., TSZ, 53 / KSA 4, 93, or KSA 13, 21.

Yet despite the close connection between power, pleasure and pain, Nietzsche also rejects this kind of reading. In his own words, ‘hedonism, pessimism, utilitarianism, eudaemonism’, insofar as they are all ‘ways of thinking that measure the value of things according to pleasure and pain’, are focused on ‘secondary states and matters. They are all foreground ways of thinking and naïvetés, and nobody who is conscious of both *formative* powers [*gestaltende Kräfte*] and an artist’s conscience will fail to regard them with scorn as well as pity.’⁸ The will to power is thus something other than the ‘will to pleasure’, or the ‘will to self-advantage’, or the ‘will to self-preservation’. These are all symptoms or consequences, but not yet the essence of Nietzsche’s diagnosis.

What, then, does one will when one *wills power*? A definitive answer to this question is not easy to find in Nietzsche’s writings, not because he forgot to tackle such an important issue, but because his main aim is precisely to overcome the philosophical prejudice underlying this kind of formulation. As he repeatedly argues, to ask about the specific *end* of the will to power is to look for one of the ‘*superfluous teleological principles*’⁹ he sought to eliminate. Whereas scientific and moral modes of valuation are grounded in a finalist approach to human life—whether focused on logical coherence, self-preservation, selflessness or self-satisfaction—a ‘philosophy of the future’ must look beyond this cognitive scheme. Power, like the ‘true’ or the ‘good’, is not something fixed, to be craved or pursued, but ‘something *to be created*, which designates a *process*, or what is more a will to overpower [*Überwältigung*] that has in itself no end.’ It is ‘a *processus in infinitum*, an *active determining*, not a becoming-conscious of something [that is] fixed and determined “in itself.”’¹⁰

The will to power is not a means to an end, but an end in itself. It is something like a pure physiological drive, a primal overflow of energy that requires no aim or justification, but merely the occasion to be released and renewed. This will is not subordinated to, nor channelled towards, a higher truth, ideal or imperative—it is itself the primary source of meaning and agency. Each of its choices is the result of a momentary relation of forces, maintained or abandoned according to whether it favours or obstructs the release and expansion of power. In this dynamic context, fixed ideas are mere obstacles: every categorical ‘yes’ or ‘no’, ‘true’ or ‘untrue’, ‘good’ or ‘evil’ is a sign of rigidity, a limitation of the endless array of possibilities open to the human will.

Nietzsche’s shift from a teleological to a physiological mode of valuation is made especially clear in a central passage from *On the Genealogy of Morality*, where he criticizes most modern accounts of the origin of legal institutions, social customs and religious rites. Once again, he compares all of these processes to the growth of a living organism, in which the notions of purpose or *telos* play a secondary and largely accidental role. What determines the evolution and growth of living things, Nietzsche argues, is their will to power and domination. Meaning and valuation are the cognitive translations of this primary will, and not the other way around:

⁸BGE, 116 / KSA 5, 160. Translation modified. Cf. KSA 13, 30 and 33ff.

⁹BGE, 15 / KSA 5, 27f.

¹⁰KSA 12, 385

Every purpose and usefulness is a mere *sign* that a will to power has achieved mastery over something less powerful, and has impressed upon it its own idea [*Sinn*] of a use function; and the whole history of a 'thing', an organ, a tradition can to this extent be a continuous chain of signs, continually revealing new interpretations and adaptations, the causes of which need not be connected even amongst themselves, but rather sometimes just follow and replace one another at random.¹¹

The meaning and ends favoured over time by a given organism, person or tradition are the records of different relations of power, domination, subjugation and resistance. Like the rings on a tree trunk, or the lines on a seismogram, they document the different forms taken on by the will to power, as well as their degree of success.

The form is fluid, but the 'meaning' [*Sinn*] even more so... It is no different inside any individual organism: every time the whole grows appreciably, the 'meaning' of the individual organs also shifts. . . . True *progressus* [always appears] in the form of a will and way to *greater power* and always [emerges] victorious at the cost of countless smaller powers. The amount of 'progress' can actually be measured according to how much has had to be sacrificed to it.¹²

This physiological mode of valuation is Nietzsche's alternative to the ingrained finalism of science and morality. In his eyes, once philosophy and psychology manage to overcome their usual prejudices, they will be free to embrace a new methodology and develop a '*doctrine of the evolution of the will to power*.'¹³

This novel task, as defined and attempted throughout Nietzsche's writings, can once again be divided into a descriptive and a prescriptive moment. The first one corresponds to Nietzsche's recurring claim that human life *in its entirety*—including every form of cognition, volition and agency—is guided by a will to power. This claim also extends, *a fortiori*, to those modes of cognition or valuation that result in a *decrease* of power, or in an outright negation of life. Indeed, insofar as these examples pose the greatest challenge to Nietzsche's physiological strategy, they are the ones he returns to more often.

Firstly, Nietzsche points out that the *epistemic* bias of philosophers and scientists is nothing other than a peculiar form of will to power, typical of overeducated beings. Like every other animal, the '*bête philosophe*' strives instinctively 'for an optimum of favourable conditions in which to fully release its power and achieve its maximum of power-sensation'.¹⁴ Its real goal is not truth or science, not even happiness or well-being. These are secondary aims, superposed to a primary drive 'to power, to action, to the mightiest deeds.'¹⁵ What distinguishes philosophers from other animals is thus the devious form usually assumed by their will to power: instead of a direct display of health and strength, 'philosophical power' is achieved by means of silence, seclusion, chastity, an abnormal interest in logic and coherence, a penchant for 'big questions' and 'essential problems'.

¹¹ GM, 51 / KSA 5, 314

¹² GM, 51f. / KSA 5, 315. Translation modified.

¹³ BGE, 23 / KSA 5, 38

¹⁴ GM, 76 / KSA 5, 350

¹⁵ Ibid.

Secondly, Nietzsche argues that the *moral* bias of priests and theologians is likewise a peculiar form of will to power, usually found in meek and unhealthy beings. In this case, his critique is less about the naïvety of this attitude and more about its hypocrisy. Like philosophers, priests are primarily driven by a will to increase and expand their power, which means that their life is not really about virtue, selflessness or compassion: these ‘noble’ values are either the products of self-delusion or deliberate ruses designed to conceal their real foundation. Physiologically, the *bête morale* differs from other animals by its natural weakness, which translates into the need for negative means of self-affirmation. Its struggle against life and health is nonetheless a struggle for power, and its situation ‘is therefore the precise opposite of what the worshippers of this ideal hold [*meinen*].’¹⁶ The priest and the theologian are both guided by a will to power—yet this power is secured and strengthened by the exaltation of impotence.

Finally, nihilism itself, as the corollary of the two previous tendencies, is ultimately but a cry for domination. The final stage of human degeneration, it exhibits the will to power in its most negative or perverted form: a will born out of absolute exhaustion, whose preferred stimulant is the negation of all stimulants; a will at war against itself, fuelled by the madness of its own destruction. Even here, power is still the operating principle: this suicidal effort ‘is and remains a will! . . . man still prefers to *will nothingness*, than *not* will. . .’.¹⁷

But if the epistemic, moral and nihilistic standpoints are indeed about power and domination, their actual results are not: they lead to slavish and self-diminishing modes of being, which imply an overall decrease in health, power and freedom. Nietzsche’s reaction to this outcome leads to the *prescriptive* moment of his philosophy, focused on the actual transformation of human existence. Once the will to power is acknowledged as the founding principle of human life, philosophy must move beyond a mere critique of existing values. The will to power must be redirected towards a truly *empowering* form, which entails both a return to the primitive values of ancient aristocracy and the ‘crossing over’ to a new and radically different world.

12.2 Return to Callicles

Looking back on the genealogical path outlined in the previous chapters, Nietzsche’s approach to the ‘problem of valuation’ can be interpreted as a response to an ancient philosophical debate, brought to the fore in Plato’s dialogues: namely, the debate regarding the definition of the greatest good or the best way of life available to human beings. In the introduction to this book and in the different sections dedicated

¹⁶GM, 88 / KSA 5, 366. Translation modified. Carol Diethe translates *meinen* as ‘imagine’, but this solution fails to capture the ambiguity of Nietzsche’s formulation: priests and moralists *hold* their ideal to be something different from what it actually is—whether out of naïvety or plain dishonesty.

¹⁷GM, 120 / KSA 5, 412

to Plato's philosophy, I called attention to the hierarchical structure of Plato's conception of human cognition. I argued, with Plato, that all of our ideas and judgements are continually shaped by a self-interested assessment of the different situations we are faced with and subordinated, whether consciously or otherwise, to what is perceived, at each given moment, as the greatest good or the best course of action available. I pointed out, furthermore, that this continual assessment is never confined to a specific event or situation, but is always related to our whole existence. All of our ideas and judgements are ultimately subordinated to a global diagnosis of what is at stake in our life—to what matters most in it, to what is essential and what is secondary, to what is worth pursuing and what is not.

This general theme is explicitly dealt with in dialogues like the *Republic*, the *Phaedo*, the *Philebus* or the *Gorgias*. In these texts, Socrates and his interlocutors are directly concerned with determining 'what course of life is best.'¹⁸ The debate usually takes the form of a dialectical contest between different values and world-views. Wisdom, pleasure or political domination are put forth as possible versions of the greatest good, and the different lives they prescribe—i.e. philosophy, hedonism or rhetoric—are likewise compared and discussed. In each case, the procedure amounts to the attempted translation of a formal principle, τὸ μέγιστον ἀγαθόν, into a concrete and consistent mode of valuation.

We find something very similar in Nietzsche's writings, where 'future philosophers' are assigned the task of determining 'the rank order of values.'¹⁹ Once again, human cognition is conceived as a global hierarchical system, whose inner composition must be brought to light. Moreover, Nietzsche also calls attention to human beings' ignorance of their own valuation system: usually, their choices and actions are put down to a more or less immediate practical concern, wrongly perceived as an end in itself; alternatively, whenever science or religion step in, thoughts and actions are led back to more general or abstract ideals. But in all of these cases the higher term of the hierarchy remains out of sight. For Nietzsche, the whole system is about *power*: viz. about the *augmentation* of the 'feeling of power' that determines and motivates every single judgement or decision. His aim is not only to bring to light the real structure of human beings' system of valuation—that is, to expose every human intention as a more or less direct instantiation of the 'will to power'—but also to align human aims and actions with this newly discovered drive—that is, to inaugurate a mode of existence that is no longer unconsciously, but *explicitly* focused on the pursuit and accumulation of power.

Yet the formal affinity between Plato's and Nietzsche's philosophical projects is even more profound. Nietzsche's proposed transition from a scientific and moral philosophy to a philosophy of power is strongly evocative of the debate held throughout the *Gorgias*, already discussed earlier. In particular, Nietzsche's views

¹⁸ *Go* 500c4

¹⁹ GM, 34 / KSA 5, 289

on power and domination are strikingly similar to those levelled by Callicles against Socrates. So much so, indeed, that Callicles and Nietzsche can almost be regarded as philosophical ‘kindred spirits’, united by their anti-Platonic views.²⁰ Their doctrines coincide on a number of crucial points, which may be summed up as follows:

1. Firstly, like Nietzsche, Callicles bases his entire argument on the sophistical distinction between the laws of nature, which regulate human life, and the laws of custom or convention, which determine what is socially acceptable.²¹ He calls attention to the contradiction between human beings’ natural drives, responsible for the growth and evolution of humankind, and the norms which have come to be endorsed in modern, ‘civilized’ societies. Furthermore, anticipating Nietzsche’s standpoint, Callicles illustrates this distinction with a genealogical account of human nature, designed to contrast the power and freedom of primitive individuals with the impotence and slavishness of modern individuals.
2. Secondly, like Nietzsche, Callicles argues that the laws of custom, responsible for the increasing domestication of humankind’s primitive instincts, are due to a majority of weaker individuals intent on protecting themselves from the danger posed by the strong. Their aim is to ‘terrorize the stronger sort of folk who are able to get an advantage, and to prevent them from getting one over *them*’. In order to do so, ‘they tell them that such aggrandizement is foul and unjust, and that wrongdoing is just this endeavour to get the advantage of one’s neighbours.’²² In Nietzschean terms, they replace nature’s *Herrenmoral* with an artificial *Sklavenmoral*, through which the aristocratic values of primitive individuals are reviled and reversed: what was good according to nature (κατὰ φύσιν)—i.e. strength, virility, strife, domination—becomes foul (αἰσχρόν) according to custom (κατὰ νόμον), and therefore evil (κακόν); conversely, what was evil according to nature—i.e. meekness, reasonableness, peace, restraint—becomes socially accepted (καλόν), and therefore good (ἀγαθόν).

Interestingly, Callicles illustrates his argument with an image that is also strikingly Nietzschean: namely, the comparison between ‘the best and strongest men’ with ‘young lions’ tamed from their infancy and ‘told the while they must

²⁰This closeness is both striking and enigmatic inasmuch as Nietzsche, who knew Plato’s work very well, does not refer once to Callicles, and only twice, in passing, to Plato’s *Gorgias* (KSA 1, 790 and KSA 9, 174). See Menzel, *Kallikles*, 80f.: ‘Wenn daher Kallikles auch von Nietzsche nicht ausdrücklich genannt wird, so darf bei der weitgehenden, oft wörtlichen Übereinstimmung kein Zweifel darüber bestehen, daß eine Einwirkung stattgefunden hat. Sie kann ja auch eine unbewußte sein. Die Eindrücke, welche der junge klassische Philologe aus der Lektüre Platons erhalten hatte, sind haften geblieben, soweit sie mit der sich entfaltenden Gedankenwelt des Philosophen Nietzsche Beziehungen besaßen.’ On the link between Nietzsche and Callicles, see also Dodds’ commentary on the *Gorgias*, 386–391; Kloch-Kornitz, ‘Der Gorgias Platons und die Philosophie Friedrich Nietzsches’; Leyra, ‘Calicles y Nietzsche’; and Urstad, ‘Nietzsche and Callicles on Happiness, Pleasure, and Power’.

²¹Regarding the translation of νόμος as ‘custom’ or ‘convention’, see Sect. 3.3, fn. 46.

²²Go 483c

have but their equal share, and that this is what is fair and just.’²³ The point, again, is that the physiological definition of power is incompatible with the moral or political definition of power, and that the latter is conditional on the artificial repression of natural qualities and instincts. It is a power *contra natura*, not only because it is *unnatural*, but also because it is obtained and maintained through an actual struggle *against nature*.

3. In addition, although Nietzsche’s and Callicles’ critique is directed at the *moral* subversion of natural values, and consists therefore in an attack on conventional justice, contractarianism and democracy, it also extends to the *philosophical* or *scientific* subversion of natural values. Both of Socrates’ opponents place the immediate power of the will above the intellectual concerns of epistemic reason, and both regard Socrates’ philosophical fixation with truth as a symptom of slavishness. Callicles is quite explicit in his warnings against the ridicule of philosophizing far on into life, instead of minding those things ‘that ought to be familiar to the man who would be a thorough gentleman and make a good figure in the world.’²⁴ On his view, philosophy represents, furthermore, a source of danger, inasmuch as the minute examination of one’s thoughts and intentions proposed by Socrates may prove paralysing and render one unable or even unwilling to defend one’s life, honour and practical interests.
4. Finally, both Callicles and Nietzsche advocate a complete reform of this state of affairs. The hypocrisy of modern values must be acknowledged as such and replaced by a new mode of valuation, in which power and strength are granted their original protagonism. Once again, the freedom provided by this transformation is described by Callicles in terms that anticipate Nietzsche’s writings: in contrast to the image of the young lions forced to hold back their strength, he envisages the emergence of a slave capable of ‘shaking all that he was taught’, ‘breaking his bonds’ and rising to the role of master.²⁵

The debate between Socrates and Callicles is centred on the notion of *power* (δυναστεία)—and, more specifically, on the determination of *what* power is, of *who* can be said to enjoy it and of *why* it is coveted in the first place. As usual, Socrates points out that none of these questions is simple or self-evident. His critique highlights a series of ambiguities that will reappear, in very similar ways, in Nietzsche’s doctrine of the will to power. Like Callicles’ law of nature (νόμος τῆς φύσεως), Nietzsche’s *Herrenmoral* is grounded in a constellation of notions whose definition is open to different and often contradictory interpretations.

²³ *Go* 483e–484a. See, for ex., *GS*, 177 / *KSA* 3, 548, where the lion is evoked by Nietzsche as a measure of personal strength; *TSZ*, 17, 129 / *KSA* 4, 30, 205, where the lion is associated with freedom and strength; or *GM*, 63 / *KSA* 5, 334, where primitive Greek men are dubbed ‘marvellous, lion-hearted children.’

²⁴ *Go* 484c9–d2

²⁵ *Go* 484a3–4

The course of the argument was already outlined above: when asked about the exact nature of the power held by superior beings, Callicles defines it as a form of strength, then wisdom, then courage and manliness. However, none of these definitions succeeds in capturing the exact phenomenon he has in mind, and Socrates' refutations lead him to settle on a fourth and final definition: power as *πλεονεξία*, that is, as the ability to 'let [one's] desires be as strong as possible and not chasten them', but rather 'minister to them when they are at their height, by reason of [one's] manliness and intelligence.'²⁶ This ability, he claims, is precisely what impotent individuals lack and the reason why they praise temperance instead of enjoyment, justice instead of force, restraint instead of freedom.

Again, the similarities with Nietzsche's position are obvious. Callicles' conviction that justice and temperance are born out of impotence and resentment anticipates Nietzsche's insistence on the 'morality of *ressentiment*', which he associates with Platonism, Judaism, Christianity and democracy. Moreover, as noted by Adolf Menzel, Callicles' notion of *πλεονεξία* is directly translated by Nietzsche as the 'will to wanting-to-have-more' (*der Wille zum Mehr-haben-wollen*), which amounts to a variation of the will to power.²⁷ Finally, and crucially, Callicles' difficulty in defining the exact nature of the power wielded by superior beings also seems to anticipate Nietzsche's unwillingness to reduce the will to power to a specific end or ideal. For Callicles, as for Nietzsche, power is not merely about strength, wisdom or courage, as though powerful individuals would simply have to submit to these values and pursue them in the best or most effective way possible. *Πλεονεξία* seems to refer, rather, to power or volition in a more basic or primordial sense: not a movement oriented towards a specific end, but a movement which is itself its own end; not a matter of achieving this or that particular goal, but the ability to renew and intensify a physiological drive for self-affirmation and self-enhancement. This idea is made clearer from 491d onwards, when Callicles condemns the ideals of personal autarky and self-mastery (*ἄρχειν ἑαυτοῦ, ἐγκράτεια ἑαυτοῦ*) as the life-goals of slaves and simpletons, and compares the alleged happiness of 'those who want nothing' to the happiness of corpses or stones. Whereas 'a man who has taken his fill can have no pleasure any more, . . . a pleasant life consists rather in the largest possible amount of inflow.'²⁸

Callicles is not focused on satisfaction as a *nomen rei actae*, but as a *nomen actionis*. He puts the emphasis not on the state being sated, but on the very movement of satisfaction. Insofar as his aim is to perpetuate and intensify a feeling of gratification, his standpoint can be interpreted as a conventional form of hedonism. Alternatively, and more interestingly, his standpoint can also be read as a preliminary attempt at the kind of 'transvaluation' championed more than two thousand years later by Nietzsche. In this light, the remainder of the *Gorgias* may likewise be read as a challenge to the cogency of Nietzsche's project and to the possibility of a radical shift from a cognitive outlook grounded in truth and morality to a *metacognitive* outlook grounded in a pure, all-encompassing 'will to power.'

²⁶ *Go* 491e8–492a2

²⁷ KSA 13, 251. Cf. KSA 9, 284 and KSA 11, 586.

²⁸ *Go* 494a6–b2

Unlike ordinary human beings, Nietzsche's *Übermenschen* must be able to create their own values, and to do so they must liberate themselves from the normative framework of truth and morality. In the *Gorgias*, Socrates' final debate with Callicles is centred precisely on whether this liberation is possible. Whilst Nietzsche aims to move beyond the oppositions truth/untruth and good/evil, Socrates reaffirms his doctrine of *δοξάζειν*, anchored in the *transcendental* value of these two oppositions. As in many other dialogues, Plato's views amount to the restatement of a global diagnosis of human cognition based on two major observations: namely a) that every human belief is necessarily supported both by a conviction of truthfulness and by a complex web of 'silent' beliefs also supported by a similar conviction; and b) that this doxastic web is not linear, but hierarchical, and in fact *universally* oriented towards the pursuit of what is perceived to be, consciously or otherwise, the greatest good or the best course of life available. In short, truth and goodness, as well as untruth and evil, are the formal standards underlying the adoption of every human belief, choice or course of action. They can be filled with different contents, or given different names—but never truly eliminated.

If the good is indeed 'the end of all our actions; if 'it is for its sake that all other things [are] done',²⁹ then every kind of life—be it the life praised by Callicles (and Nietzsche) or the one praised by Socrates—represents a *version* of the best possible life, adopted and pursued for this very reason. As all versions, however, it may be *true*, and actually beneficial, or *illusory*, and potentially harmful. It is upon this disjunction that Plato grounds his entire attack on Callicles (and Nietzsche): if every life is doomed to be a version of a true and good life, the determination of its value is ultimately dependent on whether one can actually demonstrate, rather than merely assert, that it is in fact so. In other words, an accurate mode of valuation, capable of ensuring the attainment of power, strength, wisdom or any other quality deemed beneficial or self-enhancing, is necessarily dependent on an epistemic outlook on reality—which can only be attained, if at all, through the practice of philosophy.

For Nietzsche, on the contrary, everything hinges on the possibility of refuting this line of argument and showing that our attachment to the good and the true is not as natural or as essential. As seen earlier with regard to the 'will to truth', the assessment of whether and to what extent this move is possible depends on how radically we choose to interpret Nietzsche's position. If, on the one hand, the 'transvaluation of all values' is taken to mean that human beings' standard of valuation is no longer epistemic or moral, but physiological, this may simply amount to claiming that what is usually deemed good and true will cease to be so: truthfulness and goodness will henceforth be determined by one's virility, rank, strength, health, etc., and hence applied to things which were previously deemed evil or untrue. The main tenets of truth and morality will no longer be dictated by slavish or pious beings, but by strong and virile ones. Yet although this version of the argument does accord with some of Nietzsche's formulations, it is still fairly superficial, as it affects only the *content*, and not yet the *form* of valuation. The oppositions truth/untruth and good/evil have not been overcome, but merely translated in different ways.

²⁹ *Go* 499e7–500a2

This kind of reading is behind most historical and political appropriations of Nietzsche. Indeed, many of the misunderstandings surrounding the notions of ‘over-humanity’, ‘power’ or ‘nobility’ stem from a superficial interpretation of Nietzsche’s philosophical project, subordinated to very concrete political ends. In these cases, the ‘transvaluation of all values’ has been rather shy: the overcoming of truth and morality was motivated by very specific, albeit controversial versions of the good and the true; ‘new values’, just like old ones, were oriented towards a specific identification of the greatest good, or the best possible life.

A second, more radical interpretation of Nietzsche’s project consists in the attempt to overcome the attachment to the true and the good *altogether*. Only this version is faithful to Nietzsche’s explicit intention of moving beyond teleology and purposiveness *as such*, and to free humankind from the slavishness of the ‘will to truth’ and the ‘will to the good’. But here the stakes are of course much higher: if every human judgement or decision entails, if only unconsciously or momentarily, the validation of a specific version of things (viz. a specific account of who one is, of where one stands, of what the world looks like, of what is at stake in a given situation, and so on), a life no longer supported by such validation—or, which amounts to the same thing, a life where such version were lived as a mere version, *and nothing more*—would be stripped of its most basic practical references and faced with the possibility of a complete existential paralysis.

Zarathustra voices this contradiction in the following terms:

That I must be struggle and becoming and purpose and the contradiction of purposes—alas, whoever guesses my will guesses also on what *crooked* paths it must walk! Whatever I may create and however I may love it—soon I must oppose it and my love, thus my will wants it.³⁰

In a new, ‘over-human’ life, consciousness is faced with the same kind of abyss found at the edge of Plato’s critique of human cognition: the pivotal moment when all human beliefs are perceived as mere beliefs, and thereby rejected. From that moment on, what lays in store for consciousness can only be guessed. What would life be like beyond the realm of belief? What form would human existence take? How free and powerful would one really be? For Nietzsche, as for Plato, these questions cannot be answered beforehand. They can only emerge as the result of a radical ontological transformation, whereby one’s identity, the identity of others and the meaning of the whole world would change into something else.

³⁰TSZ, 89f. / KSA 4, 148

Chapter 13

Conclusions



Abstract To conclude, I argue that neither of the philosophical projects analysed succeeds in offering a definitive remedy for the cognitive imprisonment implicit in human life. In Hegel's case, the success of the phenomenological progression is compromised by a series of formal conditions that are ultimately too ambitious to be fully met. In Nietzsche's case, a complete emancipation from the 'will to truth' seems to entail a rejection of each and every form of normativity, which appears to contradict the very structure of human existence. All things considered, it is Plato's reluctance in putting forth a final solution, either positive or negative, that best captures the problem's unique complexity. Although his criticism inaugurates a long and fruitful debate, the question he raises is left unanswered, and probably bound to remain so.

Keywords Freedom · Power · Knowledge · Truth · Plato · Hegel · Nietzsche

The itinerary followed in this book led from the ancient Greek world to the dawn of postmodernity. Starting with Socrates and Plato, I looked at some of the key moments of a long and complex philosophical debate, with concrete implications in every dimension of human life. What emerged from this survey, however, was not a clear or straightforward progression, leading to a clear or straightforward result. The examination of Plato's, Hegel's and Nietzsche's views on freedom and the pursuit of knowledge revealed a vast web of possibilities, irreducible to a single set of conclusions.

The three projects we have considered are all grounded in the same general criticism. In all of them, our usual cognitive standpoint is shown to be unreliable and in need of transformation, and in all of them this transformation is construed in a negative way, as the release from ignorance and self-deceit. According to Plato, Hegel and Nietzsche, only by eliminating the unwarranted assumptions that shape our usual understanding of reality can we hope to lead free and meaningful lives. But the nature and scope of this operation are envisaged in very different ways.

What emerged from the comparison of these three projects was a wide and complex labyrinth, where the relationship between freedom and truth can be viewed from three different angles. Each project set up a global philosophical map, as it were, where the other two count as regional locations, to be either rejected or incorporated in a wider philosophical narrative.

From Plato's perspective, Hegel's and Nietzsche's standpoints amount to sophisticated doxastic constructs, partly anticipated and refuted in the dialogues. In Hegel's case, Plato would endorse the methodical scepticism employed throughout the *Phenomenology*, but would hardly have shared Hegel's teleological approach to dialectics or the possibility of a final, all-embracing *Aufhebung*. In Nietzsche's case, Plato's reservations about the possibility of a metacognitive outlook on reality are clearly brought out in Socrates' response to Callicles: since the notion of truth is inherent to the very possibility of human life, all efforts to overcome it must be regarded with suspicion.

From Hegel's perspective, the Platonic and Nietzschean standpoints correspond to intermediate stages on the road to science and truth. In Plato's case, the dialectical exercises enacted in the dialogues are necessary, but still based on an abstract form of negation, and hence devoid of the self-movement Hegel deems essential for a truly rational philosophical enquiry. In Nietzsche's case, the direct affirmation of power echoes the master's standpoint criticized and overcome in the master-slave dialectic. Once again, this attitude is necessary but still abstract and in need of sublation.

Finally, from Nietzsche's perspective, Plato and Hegel are the key proponents of the cognitive tendency that has come to dominate and enslave human consciousness. While Plato turned the 'will to truth' into a cultural and philosophical ideal, Hegel raised this ideal to unprecedented heights. In both cases, cognition and truth were uncritically accepted as basic standards of valuation, limiting the endless diversity of human life.

On the whole, these differences are largely irreconcilable. Yet they also speak to an important similarity between Hegel's and Nietzsche's perspectives. While Plato's criticism is ultimately inconclusive, casting serious doubts on the possibility of an actual cognitive liberation, Hegel and Nietzsche aim to solve the challenge raised in the dialogues and set humankind free. But this unique ambition, which differentiates their projects from all previous forms of criticism, is also their main weakness.

For Plato, freedom entails a complete cognitive purge. However, the latter hinges on a series of requirements that are anything but self-evident. To overcome our usual pretence of knowledge, we must be able not only to track down all of the assumptions that shape our perception of reality, but also to cancel all of their effects. Moreover, for this operation to be successful, we must be able to guarantee that all assumptions have indeed been identified and cancelled. If this critical survey is not thorough enough, the release from ignorance might simply amount to a new form of delusion, grounded in a subsisting assumption or set of assumptions; worse still, since the latter are characterized precisely by their appearance of truth, there seems to be no way of guaranteeing that ignorance has indeed been overcome.

To solve this problem, Hegel turns Plato's scattered criticism into a systematic, self-moving criticism, where every stage is part of a wider organic whole. But this approach is no less uncertain. For the success of Hegel's enterprise also rests on an ambitious set of requirements, whose fulfilment cannot be guaranteed. To begin with, it hinges on there being, in fact, a finite series of stages leading from ignorance to truth. If this path is not finite, nor dividable into specific stages, the dialectical process may go on indefinitely. Moreover, for Hegel's project to be successful, its different stages must exhaust the realm of negativity and untruth. If anything important is left out, there will be no way of knowing whether the progression's final stage is really the ending, or simply a new set of presuppositions in disguise. Finally, the success of Hegel's project is also conditional on the ability to retrace and interpret its various steps. To ignore or anticipate any of the progression's stages may amount to compromising its entire meaning.

In comparison, Nietzsche's strategy appears much simpler. Instead of detailing all the assumptions that shape human cognition, his goal is to dislodge the one assumption that supports the entire cognitive edifice—the idea of truth, which Plato and Hegel left untouched. But while it is relatively easy to criticize our concern for truth as a philosophical ideal, or as a conscious life goal, it is far more difficult to relinquish truth altogether. For in its primary form, the 'will to truth' is nothing other than the need to acknowledge, at each given moment, a specific version of reality (viz. a specific version of who we are, where we stand, what the world looks like, etc.), and this need is inseparable from life itself. Accordingly, Nietzsche's attempts to dismantle human cognition are all issued from within the cognitive realm. When praising power, or strength, or vitality, he is really only endorsing another version of how things are, or should be.

Given these difficulties, neither Hegel nor Nietzsche are able to solve Plato's predicament. Though their methods succeed in questioning many of the unwarranted assumptions that guide human life, they are still haunted by the possibility of a basic presuppositive residue, immune to philosophy's critical charges. And therefore, looking back on our whole investigation, the two main challenges raised in the beginning of this book are left unfulfilled. Regarding the possibility of a truly educated outlook on reality, our enquiry has led to a negative result. Despite Hegel's and Nietzsche's efforts, we are still at odds with a standpoint exposed to ignorance and self-deceit. Consequently, regarding the possibility of a truly free outlook on reality, our enquiry has also led to a negative result. We are still at odds with an imprisoned standpoint, caught up in its own contradictions.

As was pointed out in the introduction, freedom can be defined in different ways and depend on very different requirements, including one's strength, wealth, political autonomy or good fortune. But apart from these practical criteria, freedom also entails a cognitive dimension, which is no less important. In order to be free, we must be capable of choosing and acting according to our interests, and this capacity requires the ability to make sense of reality and to determine what is at stake in each practical situation. In the absence of this ability, we may end up mistaking what we wish for what we really want, or choosing and acting against our actual interests.

Without the power to make informed and meaningful choices, there can be no real freedom. And yet this power is precisely what the present enquiry found lacking. If consciousness is not in possession of a clear and reliable account of reality—or rather, if it is unable to vouch for the clarity and reliability of its ideas and judgements—freedom is not real, but illusory. And what is more, if the power to make informed and meaningful choices underlies every other human power, this illusion is not partial, but complete and pervasive. We may well be strong, or rich, or emancipated, or fortunate, but the freedom afforded by these qualities, if unaccompanied by a suitable understanding of reality, cannot be taken for granted. For strength, wealth or good fortune to be real and empowering, they must be recognized as such, accorded a specific value and put to a specific use. And since all of these operations are based on specific ideas and judgements, continually open to refutation, all of those qualities are prone to misinterpretation and misuse. Instead of enhancing our freedom, they may amount to new sets of shackles, fastened by our own hands.

Since human cognition is ultimately unreliable, human life is not a free life. And this damning conclusion brings us back to the starting point. Philosophy's liberating efforts, in the extreme forms envisaged by Hegel and Nietzsche, have proven problematic, and Plato's negative outlook is hereby confirmed. In light of this result, we might be tempted to follow Callicles' advice to Socrates and abandon the issue altogether. Instead of probing the foundations of our ideas and judgements, in search of a solution that remains out of reach, we might simply accept their limitations and move on. But this attitude, as Plato keeps pointing out, does not solve anything. Moreover, it can only be maintained at a price. Indeed, we may well disregard the contradictions of our usual standpoint, but these are no less real and potentially debilitating. And though we may base our entire life on instinct and common sense, our choices and actions may end up contradicting our real interests. The risk involved in this voluntary ignorance is thus much higher than might be expected: in leaving behind philosophy's riddles, we may be gambling away what we value the most—the possibility of a free and meaningful life.

Bibliography

- Adorno, T.W. 1993. *Hegel. Three Studies*, trans. S.W. NicholSEN. Cambridge, MA: MIT Press.
- Aeschylus. 2009. *Persians, Seven Against Thebes, Suppliants, Prometheus Bound*, ed. and trans. Alan H. Sommerstein. Loeb Classical Library 145. Cambridge, MA: Harvard University Press.
- Alesse, F. 2007. Alcuni esempi della relazione tra l'etica stoica e Platone. In *Platonic Stoicism, Stoic Platonism*, ed. M. Bonazzi and C. Helmig, 23–40. Leuven: Leuven University Press.
- Aristophanes. 1998. *Clouds, Wasps, Peace*, ed. and trans. J. Henderson. Loeb Classical Library 488. Cambridge, MA: Harvard University Press.
- . 2000. *Birds, Lysistrata, Women at the Thesmophoria*, ed. and trans. J. Henderson. Loeb Classical Library 179. Cambridge, MA: Harvard University Press.
- . 2002. *Frogs, Assemblywomen, Wealth*, ed. and trans. J. Henderson. Loeb Classical Library 180. Cambridge, MA: Harvard University Press.
- . 2008. *Fragments*, ed. and trans. Jeffrey Henderson. Loeb Classical Library 502. Cambridge, MA: Harvard University Press.
- Aristotle et al. 1995. *Poetics, On the Sublime, On Style*, trans. S. Halliwell et al. Loeb Classical Library 199. Cambridge, MA: Harvard University Press.
- Barney, R. 2011. Callicles and Thrasymachus. In Zalta, E.N. (ed.), *The Stanford Encyclopedia of Philosophy*, Winter 2011.
- Baudelaire, C. 1868. *Oeuvres Complètes*. Vol. 2. Paris: Michel Lévy Frères.
- Beerling, R.F. 1961. Hegel und Nietzsche. *Hegel-Studien* 1 (2): 229–246.
- Blondell, R. 2002. *The Play of Character in Plato's Dialogues*. Cambridge: Cambridge University Press.
- Blumenberg, H. 1996. *Höhlenausgänge*. Frankfurt am Main: Suhrkamp.
- Breazeale, D. 1975. The Hegel-Nietzsche Problem. *Nietzsche-Studien* 4: 146–164.
- Broze, M. 1991. Calliclès et Thrasymaque: sur la ruse et la violence chez Platon. *Revue de Philosophie Ancienne* 9 (1): 99–115.
- Camus, A. 1965. *Essais*. Paris: Gallimard.
- Carvalho, M.J. 2006. Μέθοδος e ὑπόθεσις. O problema do pressuposto na fundação platónica da filosofia. In *Método e Métodos*, ed. D. Ferrer. Coimbra: Edições Universitárias de Coimbra.
- . 2009a. *Die Aristophanesrede in Platons Symposium. Die Verfassung des Selbst*. Würzburg: Königshausen & Neumann.
- . 2009b. “Verse Actuar a Sí Mesmo”. La Commedia en la Fenomenología del Espíritu. In *Figuraciones Contemporáneas de lo Absoluto*, ed. A. Alonso and M. Jiménez-Redondo. València: Universitat de València.

- Cauchi, F. 2016. Hegel and Nietzsche on thought, freedom, and “the labour of the negative”. *Journal of European Studies* 46 (2): 110–125.
- Cicero. 1989. *Orations: Pro Milone, In Prisonem, Pro Scauro, Pro Fonteio, Pro Rabirio Postumo, Pro Marcello, Pro Ligario, Pro Rege Deiotaro*, trans. N.H. Watts. Loeb Classical 252. Cambridge, MA: Harvard University Press.
- Clark, M. 1990. *Nietzsche on Truth and Philosophy*. New York: Cambridge University Press.
- Dannhauser, W. 1974. *Nietzsche's View of Socrates*. Ithaca: Cornell University Press.
- Deleuze, G. 1962. *Nietzsche et la Philosophie*. Paris: Presses Universitaires de France.
- Diels, H., and W. Kranz, eds. 1952. *Die Fragmente der Vorsokratiker*. Berlin: Weidmann.
- Dixsaut, M. 2015. *Platon-Nietzsche: L'Autre Manière de Philosophier*. Paris: Fayard.
- Dudley, W. 2007. *Hegel, Nietzsche and Philosophy. Thinking Freedom*. Cambridge: Cambridge University Press.
- Euripides. 2009. *Fragments: Oedipus-Chrysippus. Other Fragments*, ed. and trans. C. Collard, M. Loeb Classical Library 506. Cambridge, MA: Harvard University Press.
- Fichte, J.G. 1971. *Fichtes sämtliche Werke*, ed. I.H. Fichte. Berlin: Walter de Gruyter.
- Forster, M.N. 1998. *Hegel's Idea of a Phenomenology of Spirit*. Chicago: Chicago University Press.
- Fränkel, H. 1938. A Thought Pattern in Heraclitus. *American Journal of Philology* 59: 309–337.
- Gadamer, H.G. 1980. *Hegels Dialektik. Sechs Hermeneutische Studien*. Tübingen: Mohr.
- Goethe, J. 2006. *Sämtliche Werke*, ed. K. Richter. München: btb.
- Gomperz, T. 1996. *Griechische Denker. Eine Geschichte der Antiken Philosophie in drei Bänden*. Frankfurt: Eichborn.
- Gonzalez, F. 1998. *Dialectic and Dialogue: Plato's Practice of Philosophical Inquiry*. Evanston, IL: Northwestern University Press.
- Granier, J. 1966. *Le Problème de la Vérité dans la Philosophie de Nietzsche*. Paris: Seuil.
- Greene, M. 1970. Hegel's “Unhappy Consciousness” and “Nietzsche's Slave Morality”. In *Hegel and the Philosophy of Religion*, ed. D. Christensen, 125–155. The Hague: Springer.
- Guérin, M. 1975. *Nietzsche, Socrate Heroïque*. Paris: Grasset.
- Guthrie, W.K.C. 1977. *A History of Greek Philosophy*. Cambridge: Cambridge University Press.
- Hartmann, N. 1965. *Platos Logik des Seins*. Berlin: Walter de Gruyter.
- Hegel, G.W.F. 1970a. *Phenomenology of Spirit*, trans. A.V. Miller. Oxford: Oxford University Press.
- . 1970b. *Philosophy of Nature*, trans. A.V. Miller. Oxford: Oxford University Press.
- . 1971. *Early Theological Writings*, trans. T.M. Knox. Philadelphia: University of Pennsylvania Press.
- . 1977a. *Faith and Knowledge*, trans. W. Cerf, H.S. Harris. Albany: State of New York University Press.
- . 1977b. *The Difference Between Fichte's and Schelling's System of Philosophy*, trans. W. Cerf, H.S. Harris. Albany: State of New York University Press.
- . 1986. *Werke in zwanzig Bänden*, eds. E. Moldenhauer, K. Michel. Frankfurt am Main: Suhrkamp.
- . 1995. *Lectures on the History of Philosophy* (3 vols.), trans. E.S. Haldane, F.M. Simson. Lincoln: Nebraska University Press.
- . 2007. *Philosophy of Mind*, trans. W. Wallace, A.V. Miller. Oxford: Oxford University Press.
- . 2008. *Outlines of the Philosophy of Right*, trans. T.M. Knox. Oxford: Oxford University Press.
- Hegel, G.W.F., et al. 1969. *Briefe von und an Hegel. Band 1*, ed. J. Hoffmeister. Hamburg: Felix Meiner.
- Homer. 1919. *Odyssey: Books 1–12*, trans. A.T. Murray. Loeb Classical Library 104. Cambridge, MA: Harvard University Press.
- Honneth, A. 1994. *Kampf um Anerkennung. Zur moralischen Grammatik sozialer Konflikte*. Frankfurt am Main: Suhrkamp.
- Houlgate, S. 1986. *Hegel, Nietzsche and the Criticism of Metaphysics*. Cambridge: Cambridge University Press.

- Hume, D. 2001. In *A Treatise of Human Nature*, ed. D.F. Norton and M.J. Norton. Oxford: Oxford University Press.
- Hyland, D. 1995. *Finitude and Transcendence in the Platonic Dialogues*. New York: State of New York University Press.
- Hyppolite, J. 1946. *Structure et Génèse de la Phénoménologie de l'Esprit*. Paris: Aubier.
- Inwood, M. 1992. *A Hegel Dictionary*. Oxford: Wiley-Blackwell.
- Jaeger, W. 1945. *Paideia: The Ideals of Greek Culture* (2 vols.), trans. G. Highet. Oxford: Oxford University Press.
- Kant, I. 1900. *Gesammelte Schriften*. Ed. Berlin-Brandenburgische Akademie der Wissenschaften. Berlin: Reimer/Walter de Gruyter.
- . 1998. *Critique of Pure Reason*, trans. P. Guyer, A.W. Wood. Cambridge: Cambridge University Press.
- Kaufmann, W. 2013. *Nietzsche: Philosopher, Psychologist, Antichrist*. Princeton University Press: Princeton.
- Klär, I. 1969. Die Schatten im Höhlengleichnis und die Sophisten im Homerischen Hades. *Archiv für Geschichte der Philosophie* 51: 225–259.
- Kloch-Kornitz, P. 1963. Der Gorgias Platons und die Philosophie Friedrich Nietzsches. *Zeitschrift für Philosophische Forschung* 17 (4): 586–603.
- Kofman, S. 1972. *Nietzsche et la métaphore*. Paris: Payot.
- Kojève, A. 1945. Hegel, Marx et le Christianisme. *Critique* 3 (4): 339–366.
- Kümmel, F. 1968. *Hegel und Platon zur ontologischen Begründung des Zirkels in der Erkenntnis*. Tübingen: Max Niemeyer.
- Labarrière, P. 1979. *La Phénoménologie de l'Esprit de Hegel. Introduction à une Lecture*. Paris: Aubier-Montaigne.
- Leibniz, G.W. 1890. In *Die philosophischen Schriften von Gottfried Wilhelm Leibniz*, ed. C.I. Gerhardt. Berlin: Weidmann.
- Leyra, I. 2004. Calicles y Nietzsche: Dos Formas Contrapuestas de Entender el Derecho del Más Fuerte. *Nexo. Revista de Filosofía* 2: 167–176.
- Librizzi, C. 1965. *Platone o la Ricerca della Libertà*. Padova: CEDAM.
- Livingston, D. 1998. *Philosophical Melancholy and Delirium: Hume's Pathology of Philosophy*. Chicago: University of Chicago Press.
- Long, A. 2002. *Epictetus. A Stoic and Socratic Guide to Life*. Oxford: Oxford University Press.
- Long, A., and D.N. Sedley. 1990. *The Hellenistic Philosophers*. Vol. 2. Cambridge: Cambridge University Press.
- Lossi, A. 2006. *Nietzsche und Platon. Begegnung auf dem Weg der Umdrehung des Platonismus*. Würzburg: Königshausen & Neumann.
- Löwith, K. 2015. *Von Hegel zu Nietzsche. Der revolutionäre Bruch im Denken des 19. Jahrhunderts*. Hamburg: Felix Meiner.
- Marcus Aurelius. 1916. Works, ed. and trans. C.R. Haines. *Loeb Classical Library* 58. Cambridge, MA: Harvard University Press.
- Marrou, H.-I. 1948. *Histoire de l'éducation dans l'antiquité*. Paris: Seuil.
- Menzel, A. 1922. *Kallikles: Eine Studie zur Geschichte der Lehre vom Rechte der Stärkeren*. Wien: F. Deuticke.
- Muller, R. 1997. *La Doctrine Platonicienne de la Liberté*. Paris: Vrin.
- Müller, E. 2005. *Die Griechen im Denken Nietzsches*. Berlin: Walter de Gruyter.
- Natorp, H. 1903. *Platos Ideenlehre*. Hamburg: Felix Meiner.
- Nehamas, A. 1985. *Nietzsche: Life as Literature*. Cambridge, MA: Harvard University Press.
- Nietzsche, F. 1988. In *Kritische Studienausgabe*, ed. G. Colli and M. Montinari. Berlin: Walter de Gruyter.
- . 1996. *Human, All Too Human: A Book for Free Spirits*, trans. R.J. Hollingdale. Cambridge: Cambridge University Press.
- . 1997. *Daybreak: Thoughts on the Prejudices of Morality*, trans. R.J. Hollingdale, Cambridge University Press.
- . 1999. *The Birth of Tragedy and Other Writings*, trans. R. Speirs. Cambridge: Cambridge University Press.

- . 2001. *The Gay Science: With a Prelude in German Rhymes and an Appendix of Songs*, trans. J. Nauckhoff, A. del Caro. Cambridge: Cambridge University Press.
- . 2002. *Beyond Good and Evil*, trans. J. Norman. Cambridge: Cambridge University Press.
- . 2005. *The Antichrist, Ecce Homo, Twilight of the Idols*, trans. J. Norman. Cambridge: Cambridge University Press.
- . 2006a. *On the Genealogy of Morality*, C. Diethe. Cambridge: Cambridge University Press.
- . 2006b. *Thus Spoke Zarathustra: A Book for All and None*, trans. A. del Caro. Cambridge: Cambridge University Press.
- Norton, R. 1995. *The Beautiful Soul: Aesthetic Morality in the Eighteenth Century*. Ithaca: Cornell University Press.
- Paha, B. 1992. *Die schöne Seele Hegels und die Literatur der Frühromantik*. München: Grin.
- Patten, A. 1999. *Hegel's Idea of Freedom*. Oxford: Oxford University Press.
- Philonenko, A. 1997. *Leçons Platoniciennes*. Paris: Les Belles Lettres.
- . 2001. *Lecture de la 'Phénoménologie' de Hegel. Préface – Introduction*. Paris: Vrin.
- Pindar. 1997. *Olympian Odes, Pythian Odes*, ed. and trans. W.H. Race. Loeb Classical Library 56. Cambridge, MA: Harvard University Press.
- Pippin, R. 1997. *Idealism as Modernism. Hegelian Variations*. Cambridge: Cambridge University Press.
- Plato. 1914. *Euthyphro, Apology, Crito, Phaedo, Phaedrus*, trans. H.N. Fowler. Loeb Classical Library 36. Cambridge, MA: Harvard University Press.
- . 1921. *Theaetetus, Sophist*, trans. H.N. Fowler. Loeb Classical Library 123. Cambridge, MA: Harvard University Press.
- . 1924. *Laches, Protagoras, Meno, Euthydemus*, trans. W.R.M. Lamb. Loeb Classical Library 165. Cambridge, MA: Harvard University Press.
- . 1925a. *Lysis, Symposium, Gorgias*, trans. W.R.M. Lamb. Loeb Classical Library 166. Cambridge, MA: Harvard University Press.
- . 1925b. *Statesman. Philebus. Ion*, trans. H.N. Fowler. Loeb Classical Library 164. Cambridge, MA: Harvard University Press.
- . 1926a. *Cratylus. Parmenides. Greater Hippias. Lesser Hippias*, trans. H.N. Fowler. Loeb Classical Library 167. Cambridge, MA: Harvard University Press.
- . 1926b. *Laws: Books 1–6*, trans. R.G. Bury. Loeb Classical Library 192. Cambridge, MA: Harvard University Press.
- . 1926c. *Laws: Books 7–12*, trans. R.G. Bury. Loeb Classical Library 192. Cambridge, MA: Harvard University Press.
- . 1959. *Gorgias. A Revised Text with Introduction and Commentary by E.R. Dodds*. Oxford: Oxford University Press.
- . 2013a. *Republic: Books 1–5*, ed. and trans. C. Emlyn-Jones, W. Preddy. Loeb Classical Library 237. Cambridge, MA: Harvard University Press.
- . 2013b. *Republic: Books 6–10*, ed. and trans. C. Emlyn-Jones, W. Preddy. Loeb Classical Library 237/238. Cambridge, MA: Harvard University Press.
- Plutarch. 1917. *Plutarch's Lives: Agesilaus and Pompey. Pelopidas and Marcellus*, trans. B. Perrin. Loeb Classical Library 87. Cambridge, MA: Harvard University Press.
- Proust, M. 1954. *À la Recherche du Temps Perdu*. Paris: Gallimard.
- Riedel, M. (ed.) 1990. *Hegel und die antike Dialektik*. Frankfurt am Main: Suhrkamp.
- Riginos, A.S. 1976. *Platonica. The Anecdotes Concerning the Life and Writings of Plato*. Leiden: Brill.
- Röttges, H. 1981. *Der Begriff der Methode in der Philosophie Hegels*. Meisenheim am Glan: Hain.
- Rousseau, J.-J. 1954. *Du Contrat Social, ou Principes du Droit Politique*. Paris: Garnier Frères.
- Schiller, F. 1879. *Poems of Schiller*, ed. and trans. H.D. Wireman. Philadelphia: Sherman & co.
- . 2015. *Friedrich Schiller. Poet of Freedom*, vol. 1, trans. J. Sigerson et al. New York: Schiller Institute.
- Schopenhauer, A. 1965. In *Sämtliche Werke*, ed. W. von Löhneysen. Stuttgart/Frankfurt am Main: Cotta-Insel.

- Seneca, L.A. 1917. *Epistles 1–65*, trans. R.M. Gummere. Loeb Classical Library 75. Cambridge, MA: Harvard University Press.
- Smith, S.B. 1991. *Hegel's Critique of Liberalism. Rights in Context*. Chicago: University of Chicago Press.
- Smith, R. 2018. The Ancient Quarrel and the Dream of Writing. *Journal of Philosophy of Education* 52 (4): 609–626.
- Solomon, R. 1985. *In the Spirit of Hegel*. Oxford: Oxford University Press.
- Sophocles. 1994a. *Ajax, Electra, Oedipus Tyrannus*, ed. and trans. H. Lloyd-Jones. Loeb Classical Library 20. Cambridge, MA: Harvard University Press.
- . 1994b. *Antigone. The Women of Trachis. Philoctetes. Oedipus at Colonus*, ed. and trans. H. Lloyd-Jones. Loeb Classical Library 21. Cambridge, MA: Harvard University Press.
- Statkiewicz, M. 2009. *Rhapsody of Philosophy: Dialogues with Plato in Contemporary Thought*. Penn State University Press.
- Stekeler-Weithofer, P. 2014. *Hegels Phänomenologie des Geistes. Ein Dialogischer Kommentar*. Hamburg: Felix Meiner.
- Stern, R. 2015. *Kantian Ethics. Value, Agency, and Obligation*. Oxford: Oxford University Press.
- Striker, G. 1994. Plato's Socrates and the Stoics. In *The Socratic Movement*, ed. P. Waerdt, 241–251. Ithaca: Cornell University Press.
- Taylor, C. 1975. *Hegel*. Cambridge: Cambridge University Press.
- Trisokkas, I. 2012. *Pyrrhonian Scepticism and Hegel's Theory of Judgement*. Brill: Leiden.
- Urstad, K. 2010. Nietzsche and Callicles on Happiness, Pleasure, and Power. *Kritike* 4 (2): 133–141.
- Viellard-Baron, J.-L. 1979. *Platon et l' idéalisme allemand*. Paris: Beauchesne.
- Vlastos, G. 1991. *Socrates: Ironist and Moral Philosopher*. Ithaca: Cornell University Press.
- Walsh, W. 1969. *Hegelian Ethics*. London: Macmillan.
- Wieland, W. 1982. *Platon und die Formen des Wissens*. Göttingen: Vandenhoeck & Ruprecht.
- Williams, R. 1992. *Recognition: Fichte and Hegel on the Other*. Albany: State of New York University Press.
- . 2012. *Tragedy, Recognition, and the Death of God: Studies in Hegel and Nietzsche*. Oxford: Oxford University Press.
- Wolfsdorf, D. 2008a. The Method ἐξ ὑποθέσεως at 'Meno' 86e1–87d8. *Phronesis* 53 (1): 35–64.
- . 2008b. *Trials of Reason. Plato and the Crafting of Philosophy*. Oxford: Oxford University Press.
- Wood, A. 1990. *Hegel's Ethical Thought*. Cambridge: Cambridge University Press.
- Zimmerman, R. 2005. *The Kantianism of Hegel and Nietzsche. Renovation in 19th-Century German Philosophy*. Lewiston/Queenston/Lampeter: Edwin Mellen Press.

Index

(Socrates, Plato, Hegel and Nietzsche were not included in the following list, as they are continually mentioned throughout this work.)

A

Abraham, 189
Achilles, 115
Adorno, T.W., 149
Aeschylus, 12
Aesop, 259
Alcibiades of Scambonidae, 85, 86, 119
Antigone, 12, 206, 207, 215
Apollo, 261, 264
Aquinas, T., 194
Archimedes of Syracuse, 81
Aristophanes, 11, 81, 82
Aristotle, 40, 174, 216
Augustine of Hippo, 194

B

Baudelaire, C., 218
Blumenberg, H., 40, 60, 111, 115, 238

C

Callicles of Acharnae, 43, 82–86, 90, 92–95,
111, 116, 183, 184, 199, 272, 273,
293–299, 302, 304
Camus, A., 266
Chrysippus, 20
Creon of Thebes, 206, 207, 215

D

Danton, G., 210
Darwin, C., 290

Deleuze, G., 40, 42, 279
Derrida, J., 40
Descartes, R., 163
Diderot, D., 208
Diogenes Laertius, 95
Dionysos, 261, 264

E

Engels, F., 179
Epictetus, 20, 183, 184
Eteocles, 206
Euripides, 83, 116, 262

F

Fichte, J.G., 34, 36, 169, 175, 197
Fränkel, H., 13, 98, 238, 239

G

Gadamer, H.-G., 172, 179
Glaucón, 92, 124
Goethe, J.W., 55, 57, 172, 212
Gorgias of Leontini, 23, 43, 71, 77, 82, 87–89,
91, 92, 94, 95, 116, 183, 184, 272–274,
294, 295, 297, 298

H

Habermas, J., 40
Heidegger, M., 37, 65, 232
Heraclitus of Ephesus, 13, 239

Herder, J.G., 206
 Hippias of Elis, 77
 Hobbes, T., 177, 204
 Hölderlin, F., 206
 Homer, 115, 116
 Honneth, A., 175
 Hume, D., 119–121, 127
 Husserl, E., 37, 65, 232

I

Isaiah (prophet), 194

J

Jesus Christ, 189
 Job, 77, 78, 190
 Jocasta, 8

K

Kant, I., 21, 34–36, 65, 104, 132, 140, 141, 165,
 197, 203, 210–212, 233, 242, 283
 Kierkegaard, S., 232
 Kojève, A., 175, 176

L

Leibniz, G.W., 59, 105, 151, 152, 164, 194, 283
 Löwith, K., 232
 Luther, M., 283

M

Marcus Aurelius, 20, 183
 Marx, K., 175, 179, 180, 232
 Mendelssohn, M., 36
 Moses, 189

N

Napoleon, 283
 Natorp, P., 104, 105

O

Odysseus, 85, 114, 115
 Oedipus, 7, 8, 16, 206

P

Penelope, 128
 Pericles, 85
 Pindar, 114, 115
 Plutarch, 81
 Polus of Acragas, 87, 89–92, 94, 272, 273
 Polynices, 206
 Prodicus of Ceos, 77
 Prometheus, 12
 Protagoras of Abdera, 77
 Protarchus of Athens, 51, 63
 Proust, M., 138
 Pygmalion, 250
 Pyrrho of Elis, 246

R

Robespierre, M., 210
 Rousseau, J.-J., 200, 204, 210, 212

S

Saint-Just, L.A., 210
 Schelling, F.W.J., 34–36, 227
 Schiller, F., 3–5, 206, 258
 Schopenhauer, A., 19, 36, 232, 242, 260, 261,
 284
 Seneca, L.A., 128
 Sextus Empiricus, 189
 Silenus, 260, 281
 Sophocles, 7, 8, 12, 13, 16, 83, 114, 115, 206
 Spinoza, B., 290

T

Thales of Miletus, 80, 82, 189
 Theodorus of Cyrene, 76
 Thrasymachus of Calchedon, 77, 92, 93, 284
 Tiresias, 8, 115

V

Vitruvius, 81

Z

Zarathustra, 33, 238, 240, 243, 246, 258, 299
 Zeus, 106, 114